Review Vol. 12, No. 1 / January 2025 / Optica 113

OPTICA

Computational optical imaging: on the convergence of physical and digital layers

Zhaoqiang Wang,[1] Yifan Peng,[2] Lu Fang,[3] and Liang Gao[1,*]

[1]Department of Bioengineering, University of California, Los Angeles, Los Angeles, California 90095, USA
[2]Department of Electrical & Electronic Engineering and Computer Science, The University of Hong Kong, Hong Kong SAR, China
[3]Department of Electronic Engineering, Tsinghua University, Beijing, China
*gaol@ucla.edu

Received 15 October 2024; revised 18 December 2024; accepted 19 December 2024; published 17 January 2025

Optical imaging has traditionally relied on hardware to fulfill its imaging function, producing output measures that mimic the original objects. Developed separately, digital algorithms enhance or analyze these visual representations, rather than being integral to the imaging process. The emergence of computational optical imaging has blurred the boundary between hardware and algorithm, incorporating computation in silico as an essential step in producing the final image. It provides additional degrees of freedom in system design and enables unconventional capabilities and greater efficiency. This mini-review surveys various perspectives of such interactions between physical and digital layers. It discusses the representative works where dedicated algorithms join the specialized imaging modalities or pipelines to achieve images of unprecedented quality. It also examines the converse scenarios where hardware, such as optical elements and sensors, is engineered to perform image processing, partially or fully replacing computer-based counterparts. Finally, the review highlights the emerging field of end-to-end optimization, where optics and algorithms are co-designed using differentiable models and task-specific loss functions. Together, these advancements provide an overview of the current landscape of computational optical imaging, delineating significant progress while uncovering diverse directions and potential in this rapidly evolving field.

https://doi.org/10.1364/OPTICA.544943

1. INTRODUCTION

Optical imaging was traditionally seen as the science of lenses. However, the invention of the charge-coupled device (CCD) in 1969, which enabled image data to be stored in electronic format, bridged it with digital processing. Advancements in information theory and the development of more powerful computer processors have led to the creation of numerous algorithms designed to enhance and analyze image data [1]. Post-detection processing now allows for distortion correction through geometric transformations [2] and image sharpening via deconvolution [3]. These enhancements can be achieved without the need for intricate lens optimization, shifting the responsibility for high-quality image formation from optics alone to computational algorithms as well.

This shift has given rise to the field of computational imaging, which involves the joint optimization of both optical systems (physical layer) and computational algorithms (digital layer), rather than treating them independently. In computational optical imaging (Fig. 1), digital processing is an integral part of the imaging system, rather than an afterthought. The optical system, including lenses, modulators, and other optoelectronic devices, encodes the object's properties into the wavefront. The camera sensor then captures this light, and digital algorithms act as decoders, estimating the desired information from the pixel readouts. This optical encoder–electronic decoder model implies that the wavefront at the sensor does not necessarily need to form a traditional image, as in conventional optical systems. Instead, it should convey specific information in a way that can be efficiently interpreted by digital processes. By introducing a new degree of freedom through computational processes, this framework has the potential to relax the constraints of conventional optical imaging design, achieving comparable performance with cost-effective, compact designs, or enabling unprecedented imaging capabilities.

In this mini-review, we explore the advancements in the field broadly defined as computational optical imaging (COI). The progress in computational methods has inspired unconventional optical designs, while emerging optoelectronic components are fostering optical computing to address limitations in their digital counterparts. This review aims to highlight the evolving roles and functions of physical and digital layers in COI and discusses methodologies that efficiently leverage this hybrid design for novel applications. We begin by outlining the fundamental limitations of conventional optical imaging in Section 2, emphasizing the necessity for computational solutions. In Section 3, we discuss scenarios where the physical and digital layers are co-designed to address these limitations. The presented designs are often driven by explicit mathematical models and are not optimized concurrently.

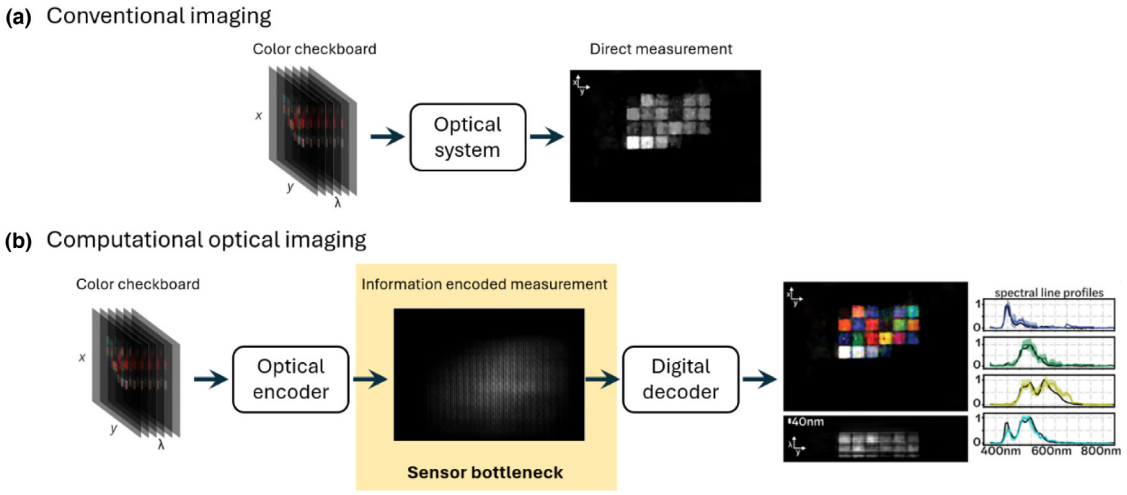

Fig. 1. Comparison of (a) conventional and (b) computational optical imaging architectures. In contrast to directly forming an image, computational optical imaging first encodes the object's information, which could be high dimensional. The captured image may look nothing like the desired final result. The digital layer then decodes this encoded data to produce task-specific results. By relaxing the constraint that the captured data should resemble the object, we can optimize the encoded data to enhance the synergy between these layers. Examples shown here are adapted from Ref. [4] Optica Publishing Group.

Section 4 then explores a data-driven approach of end-to-end optimization, where both layers are aligned through the simulation of hardware using a differentiable model, with the goal of optimizing the design of optics and algorithms under a unified task-specific loss function. Finally, Section 5 summarizes the key challenges and proposes future directions for the field.

2. BARRIERS IN CONVENTIONAL OPTICAL IMAGING

A. Physical Limitations

Despite advanced computer-aided design tools and fabrication methodologies, conventional optical systems are subject to fundamental limitations. One such limitation is the diffraction limit, which describes the smallest resolvable features by a lens. This limit is imposed by the diffraction of light wavefronts as they pass through the lens aperture, constraining the spatial resolution to approximately half the light wavelength, even with the highest-quality lens elements [5].

Another fundamental physical limit is sensitivity, as captured images are always corrupted by noise [6,7]. For example, photon shot noise, which arises from the discrete nature of photons in a light source, is an intrinsic fluctuation in the number of photons detected over a given period, resulting from the random arrival times of individual photons. This type of noise follows a Poisson distribution, where the variance is equal to the mean number of detected photons. It becomes particularly significant in low-light imaging conditions, fundamentally limiting the image signal-to-noise ratio by the number of photons received at each image pixel. This issue is further exacerbated in dynamic imaging, as the limited number of photons are spread out into temporal bins, reducing the available signal at each temporal frame even further.

B. Dimensionality Gap

While conventional optical imaging primarily captures two-dimensional (2D) trichromatic intensity images, light possesses more properties than just the spatially variant irradiance in three color channels [8]. Other light properties, such as propagation angles, wavelength, and polarization, offer many potential channels for transmitting information. However, these properties are often irreversibly lost by standard sensors that integrate over plenoptic dimensions [9].

The primary challenge in acquiring multidimensional plenoptic light data is adapting standard image sensors in the relevant spectral range to measure this information without significant hardware modifications, thus leveraging existing camera sensor technologies. A common approach to this challenge is to acquire information sequentially. For instance, in hyperspectral imaging, capturing a three-dimensional (3D) (x, y, λ) light datacube with a 2D image sensor typically involves trading a spatial axis for spectral information. The system disperses light using a prism or grating and measures a spatio-spectral (x, λ) slice per camera readout, requiring scanning along the remaining spatial dimension (y) to construct the full datacube. Similarly, in ultrafast imaging, a streak camera captures temporal information by deflecting light with a sweeping voltage, mapping photon time-of-arrival to a spatial axis and producing a spatiotemporal (x, t) slice per readout. Repeating this measurement across each y location yields an (x, y, t) event datacube.

Although these methods can provide high resolution, the scanning mechanism demands strict repeatability of the scene; the light datacube must remain unchanged during scanning. However, this condition is often compromised by object motion in hyperspectral imaging or the stochastic nature of events in ultrafast imaging.

C. Inefficient Imaging Workflow

Despite diverse applications, imaging systems mostly mimic what a single human eye perceives: capturing a sharp clear image within a two-dimensional field-of-view. Subsequently, algorithms are developed to perform high-level tasks such as object detection and classification. While this application-agnostic design is versatile, it is not necessarily the most efficient or optimal way to measure the information needed for a specific task. The limited sensor bandwidth, reflected by the number of pixels transmitted in unit time, serves as the bottleneck of the optical system, particularly in

real-time measurements [10]. Optimizing imaging hardware to make the best use of this bandwidth according to the task at hand represents an unmet need in optical imaging.

3. COMPUTATIONAL OPTICAL IMAGING: REALIZATIONS AND APPROACHES

Despite a long history of continued efforts, very few conventional optical systems can effectively tackle these bottlenecks. These challenges have spurred the recent development of computational optical imaging, which integrates innovative optical system designs with algorithms to collaboratively execute imaging tasks. Concurrently, advancements in optoelectronic components have enabled computation to occur alongside digital backends with optimized efficiency. In this section, we illustrate the interplay between physical and digital layers through several representative methods.

A. Computational Solutions to Bypass Physical Layer Limitations

1. Resolution

As discussed in Section 2, conventional optical imaging systems are inherently constrained by the diffraction limit, imposing a fundamental resolution barrier. While various hardware and photochemical methods have been explored to transcend this limitation [5,11–13], computational imaging emerges as a promising, cost-effective alternative, offering broader accessibility to the research community.

A key principle in computational super-resolution imaging is to encode the high-frequency components of an object's power spectrum into the measurement by employing angled or structured illumination, effectively shifting them towards the low-frequency range so they can be captured by the limited aperture of an imaging system. Two prominent techniques that utilize this approach are Fourier ptychography microscopy (FPM) [14–16] and structured illumination microscopy (SIM) [17,18] (Fig. 2).

Fourier ptychography microscopy: FPM allows the capture of high-resolution images using low-cost, small-aperture systems by capturing multiple images with varying illumination angles and computationally reconstructing high-resolution images across a large field-of-view (e.g., an effective NA of 1.6 using a 10 × /0.4 NA objective lens [19]).

Illuminating the object with an on-axis plane wave results in an addressable spatial frequency range that forms a circular area in the Fourier domain, with a diameter of $(1/\lambda)\cdot$NA, known as the diffraction-limited frequency bandwidth. Here, λ represents the wavelength, and NA is the numerical aperture of the imaging lens. The center of this circle aligns with the origin of the Fourier space. When angled illumination is applied, the center of the diffracted light cone shifts according to the angle of illumination, causing a linear displacement of the addressable frequencies in the Fourier domain. This allows the system to capture high-frequency components of the object that would otherwise be blocked by the aperture. By acquiring a series of images under different illumination angles, we can synthetically build up a larger effective aperture, corresponding to the sum of the illumination and objective apertures. Thus, with a small NA objective and large NA programmable illumination (e.g., [20]), super-resolution factors of 10× or greater can be achieved, albeit at the cost of requiring many images and thus long acquisition times. To reconstruct the high-resolution image, a straightforward inverse Fourier transform of the combined frequencies is performed. However, this process requires capturing the phase information as well. This is accomplished by ensuring sufficient overlap in the spatial frequency coverage of adjacent shifts in Fourier space [21] and solving an inverse optimization problem.

Because FPM recovers not only a high-resolution intensity image but also a phase map, it is frequently used for quantitative phase imaging [22–24]. Notably, FPM does not involve direct phase measurement; instead, phase information is recovered from intensity images through an iterative process, without need for reference beams or strictly coherent illumination. This capability allows FPM to bypass challenges associated with interferometry-based techniques, such as speckle noise and sensitivity to phase errors.

Traditional FPM assumes the object is thinner than the objective's depth-of-field; when dealing with thick 3D objects, tilting the illumination alters the object's power spectrum rather than merely shifting it in the Fourier domain. This effect can be incorporated in a forward model, in order to solve the 3D inverse

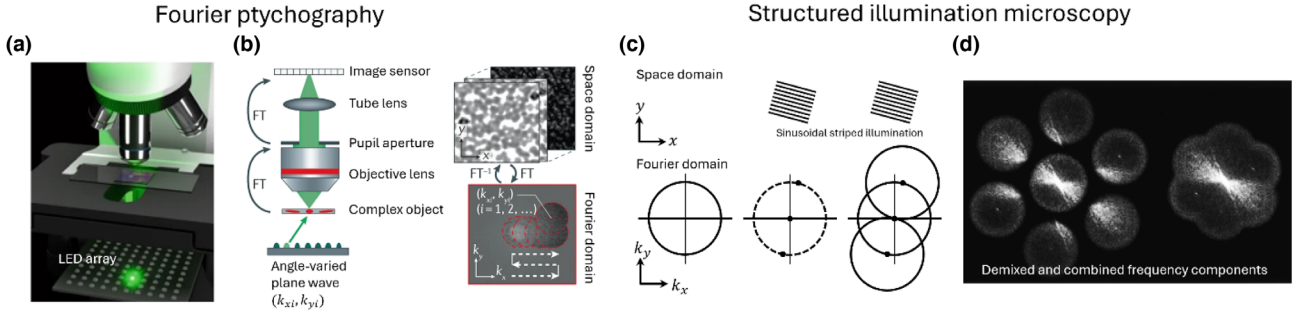

Fig. 2. (a) The setup of Fourier ptychography microscopy. (b) The principle of Fourier ptychography. By illuminating the sample with a sequence of angle-varied plane waves, the camera detects intensity images that retrieve different spectrum components in the Fourier domain through an iterative algorithm. The red dotted circles represent the support constraint by a finite pupil aperture. (a) and (b) are reproduced with permission from Springer Nature [14]. (c) The principle of structured illumination microscopy. Conventional microscopes are limited by diffraction, thus being a low-pass filter in Fourier domain (the center circle). Sinusoidal illumination pattern has three Fourier components, represented by the black dots. They shift the sample spectrum so that additional high-frequency components are detected by the microscope bandwidth. (d) Through structured illuminations of different orientations and phases, high-frequency components can be demixed from measurement and recombined into an extended spectrum. (c) and (d) are reproduced with permission from John Wiley & Sons [18].

problem and reconstruct a thick object, if the data contains sufficient diversity and redundancy [25–27]. In modeling thick objects, a multi-slice approach to treating the object like a series of thin slices, with free-space propagation between them, can further account for multiple scattering of light in the sample and may be able to recover 3D complex-field maps for highly scattering samples. To provide more diversity in the measurements, recent work further employed fixed illumination [28] and modulated light waves in the detection path [29].

Additionally, developments in light scattering models have paved the way for reflection-mode FPM [30,31], which, when combined with modulation techniques, shows promise for deep tissue imaging [32]. One limitation of FPM is its inapplicability to fluorescent samples. This is because fluorescence emission is generally isotropic and independent of illumination angles, making the angle-dependent techniques of FPM ineffective.

Structured illumination microscopy: like FPM, SIM achieves super-resolution imaging by shifting the object's power spectrum in the frequency domain. However, unlike FPM, SIM accomplishes this through spatially patterned illumination, making it applicable to fluorescent objects. In a typical SIM setup [Fig. 2(c)], the sample is illuminated with a striped pattern of a specific spatial frequency, k_1. The resulting image is a product of the object function and the illumination pattern. Due to the frequency-shifting property of the Fourier transform, the power spectrum of the image is shifted by k_1 relative to the original spectrum, allowing high frequencies beyond the imaging system's aperture to be captured. Alternatively, the illumination can be an unknown speckle pattern, which also shifts the spatial frequencies, but in a random way; with enough patterns, the super-resolved image can still be recovered computationally [33].

SIM is primarily implemented in epi-illumination (reflection) mode, where the illumination and imaging paths share the same objective lens. This constrains the illumination pattern frequency to be within the bandwidth of the objective lens for linear imaging, resulting in a maximum twofold improvement in resolution [18,34].

Employing nonlinear excitation can further enhance resolution. In this case, the nonlinear dependence of fluorescence emission on illumination power causes the effective illumination pattern to produce harmonics with spatial frequencies that are multiples of k_1. As a result, the object's power spectrum is shifted by nk_1, where n represents the nonlinear power dependence, with $n > 1$. This effectively expands the system's frequency bandwidth beyond what is achievable with linear excitation. In particular, a nonlinearity that is nonpolynomial (i.e., has an infinite Taylor series), such as in saturation excitation, can generate an infinite number of harmonics. This, in theory, could lead to infinite resolution [35]. However, practical limitations such as signal-to-noise ratio and photostability ultimately constrain the achievable resolution.

2. Photon Noise

In optical imaging, photon noise exacerbates measurement uncertainty and constrains imaging speed and sensitivity. Although hardware approaches encounter limitations in overcoming this physical barrier, comprehensive computational strategies demonstrate potential to enhance the system's resilience to photon noise through the development of data-driven denoising algorithms.

Image denoising has been a longstanding focus in computer vision, with many algorithms, such as non-local means, block-matching 3D (BM3D), and wavelet transform, developed and proven effective in general imaging tasks [36]. However, removing photon noise presents a more complex challenge due to its signal-dependent amplitude. Recent advances in deep learning, particularly through supervised and self-supervised methods, have shown great promise in outperforming conventional approaches in denoising photon-scarce images.

Supervised methods require a set of noisy input images paired with their clean counterparts. These approaches are highly effective in practice and often yield the highest-quality results. However, the necessity for paired training data can be problematic. To achieve optimal results, the training data must match the type of data being denoised. This means that ideally, new training data should be acquired for each new experiment to avoid artifacts. One way to gather such training data is by recording both low- and high-power illumination images prior to starting an experiment [37,38]. Nonetheless, in situations where obtaining ground truth data by extending the exposure time or altering illumination is infeasible, such as when the scene is dynamically changing and non-repeatable, supervised learning becomes unsuitable.

In contrast, self-supervised methods allow for training using only noisy images. Lehtinen *et al.* demonstrated that a neural net trained on pairs of measurements with independent, zero-mean noise can approximate the model derived from using ground-truth images [39]. However, in this Noise2Noise (N2N) approach, pairs of measurements must be taken on the same target signal, which are not always available. To address this limitation, methods like Noise2Void [40] and Noise2Self [41] train the denoiser directly on the noisy image itself. While the noise exhibits statistical independence, these approaches assume the true signal exhibits correlation between pixels. Training pairs are synthesized between surrounding and center pixels, or between subsets of the image. A denoising function can then be learned by feeding such data to the network.

In the following sections, we discuss several recent studies that successfully apply these strategies to photon-starved imaging, with a particular focus on fluorescence microscopy (Fig. 3).

Time-domain methods: Li *et al.* developed a self-supervised learning method called DeepCAD for denoising dynamic optical imaging data [42]. Their method leverages the temporal redundancy in video-rate imaging by treating any two consecutive frames as independent samples of the same underlying event. Consequently, image pairs composed of consecutive frames can be used for training denoising models. Furthermore, instead of training the model using a single pair of 2D frames, they decomposed the raw video into two image sequences of interleaved temporal frames. These sequences were used as the input and output data for their model, allowing them to fully exploit spatiotemporal information in the time-lapse image stack. Later, the same group developed the DeepCAD-RT model, which retained the self-supervised concept of splitting adjacent frames into inputs and corresponding targets for training a DNN but with significantly improved performance [44]. They demonstrated the capability and versatility of DeepCAD-RT in a series of photon-limited bio-imaging experiments, effectively operating beyond the shot-noise limit.

Spatial-domain methods: Zhang *et al.* developed the DeepSeMi self-supervised learning method for denoising single-frame image data. Building on the same conceptual thread

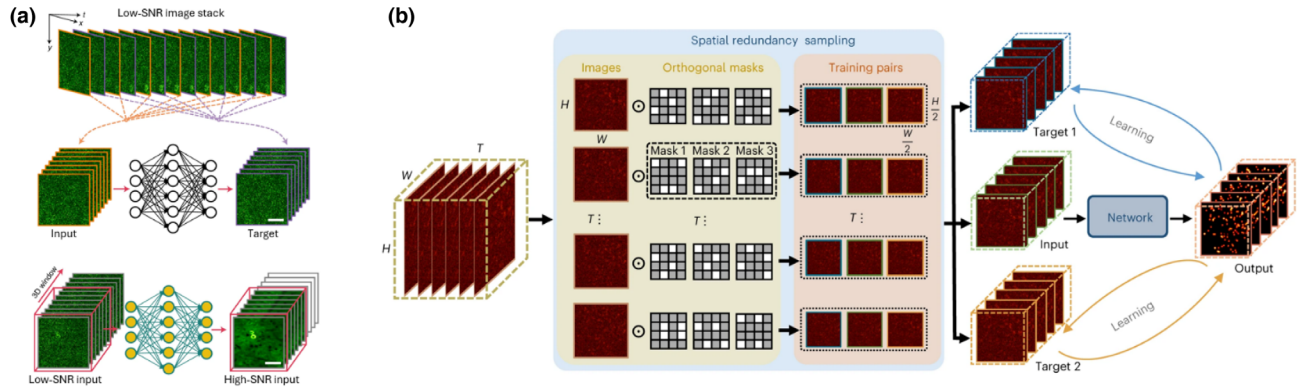

Fig. 3. Innovations in deep neural networks overcoming photon noise limitations. (a) Consecutive frames of low-SNR data are divided into two substacks for input and target volume for denoising network training [42]. (b) The original low-SNR stack is sampled by orthogonal masks, producing downsampled substacks for self-supervised training of a denoising network [43]. Figures are reproduced with permission from Springer Nature [42] and under a Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/) [43].

as Noise2Void model [40], DeepSeMi assumes that mutual mappings from neighboring pixels to a centered pixel can be established, excluding the centered pixel itself, due to local structure continuity [45]. Although these mappings can be significantly degraded under noisy conditions, the average of the degraded mappings still points to the clear pixel information. This approach ensures that the network does not simply memorize the noisy values but learns to infer the underlying clean signal based on the context provided by neighboring pixels. Due to its tolerance for noisy data, DeepSeMi can image biological dynamics under low-light conditions, significantly minimizing phototoxicity and photobleaching. This capability allows for extended observation periods, facilitating long-term studies of biological processes.

As DeepSeMi relies on the immediate neighborhood of each pixel to predict its value, its receptive field is inherently limited. This limitation restricts its ability to denoise structures that require an understanding of more global context within the image. To address this issue, Li *et al.* proposed the Spatial Redundancy Denoising Transformer (SRDTrans) [43]. This method uses orthogonal masks to decompose a single-frame image into multiple sub-images, designating one as the input and the others as targets. Additionally, they designed a lightweight spatiotemporal transformer architecture that captures long-range dependencies and high-resolution features with low computational cost. SRDTrans was demonstrated to effectively restore high-frequency information without producing oversmoothed structures or distorted light intensity traces over time in various fluorescence microscopy applications.

Lastly, in cases where redundant data cannot be provided through physical measurements, Qiao *et al.* demonstrated that such data can be generated using a Recorrupted-to-Recorrupted (R2R) scheme [46]. Their method, known as the Zero-Shot Deconvolution Network (ZS-DeconvNet), generates two noise-independent recorrupted images from the original image, which are then used as inputs and targets in network training [47]. By re-corrupting the noisy images, ZS-DeconvNet creates a more challenging training scenario, encouraging the network to develop stronger denoising capabilities as it learns to differentiate between multiple layers of noise. In addition to denoising, ZS-DeconvNet also incorporates a deconvolution layer to enhance the resolution. This dual capability allows ZS-DeconvNet to overcome both resolution and shot-noise limits within a single digital toolbox, making it a powerful solution for high-quality image reconstruction in scenarios with significant noise and resolution constraints.

3. *Dimensionality Gap*

The dimensionality gap arises from the need to capture high-dimensional plenoptic information with the limited capacity of low-dimensional photon detectors available in specific wavelength ranges. Computational optical imaging addresses this challenge by multiplexing the high-dimensional plenoptic information onto low-dimensional detectors and solving the inverse problem digitally [48]. Two established strategies for manipulating light dimensionality are computed tomography and dimension encoding (Fig. 4).

Computed tomography: tomography was traditionally used to reconstruct volumetric images by capturing 2D projections from multiple view angles. Snapshot 3D optical imaging can be achieved by multiplexing the projection information onto a 2D image sensor. For example, light field imaging shares strong similarity with limited angle tomography [25,53–56]. Its 4D ray parameterization (location + angle) can be interpreted as simultaneous acquisition of 2D sub-aperture projections from different viewpoints. With implementations such as a microlens array [55,57–59], camera array [60], and diffuser [61–63], the volumetric datacube (x, y, z) can be reconstructed from single 2D measurement in post-processing by tomographic reconstruction [54], deconvolution [56,64], and increasingly, learning-based algorithms [65–69]. Additionally, tomographic acquisition techniques have been integrated with light field detection, further reducing measurement dimensionality to 1D, thereby enabling the use of line image sensors [70–72].

Beyond volumetric imaging, the tomographic approach has been expanded to capture light information along other dimensions, such as spectrum and time. For instance, computed tomography imaging spectrometry (CTIS) is a computational spectral imaging technique capable of acquiring a 3D (x, y, λ) datacube in a single snapshot [49,73–78]. CTIS positions a computer-generated hologram disperser at the conjugate plane of the imaging system's aperture stop. Unlike a conventional diffractive grating that diffracts light along one dimension, the one in CTIS is designed to diffract light along two dimensions. The images associated with different diffraction orders undergo

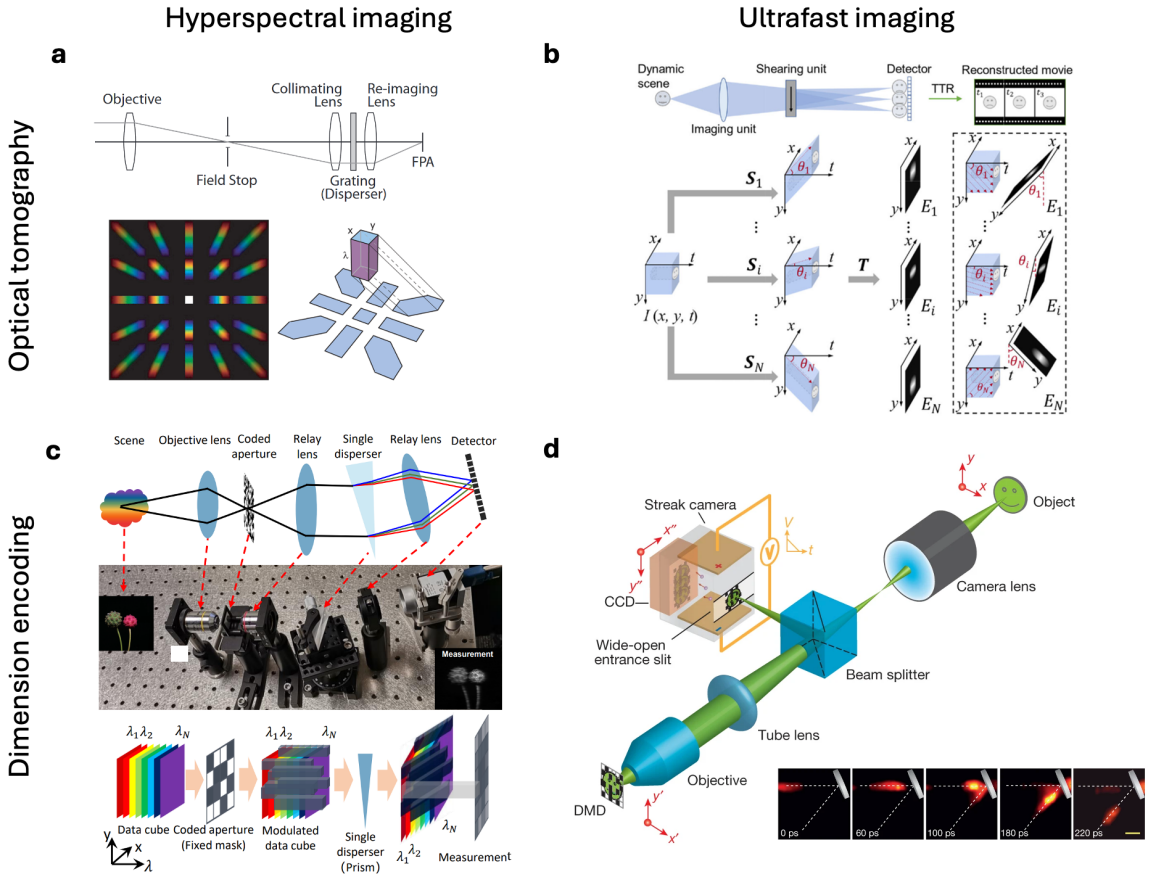

Fig. 4. Computational solutions addressing the dimensionality gap problem. (a) Computed tomography imaging spectrometry (CTIS) [49]. (b) Compressed ultrafast tomographic imaging (CUTI) [50]. (c) Single-disperser coded aperture snapshot spectral imaging (CASSI) [51]. (d) Compressed ultrafast photography (CUP) [52]. Figures are reprinted with permission from Refs. [49,50] Optical Society of America and from Refs. [51,52] Springer Nature.

varying levels of spectral dispersion along different in-plane axes. Each spectrally dispersed image can be considered a tomographic projection of the spectral datacube along a specific view angle in (x, y, λ) space, enabling the application of a standard back-projection algorithm for image reconstruction.

CTIS has been demonstrated in various imaging applications, including microscopy [79,80], astronomy [81], and ophthalmoscopy [82]. However, CTIS is inherently a limited-view instrument—each voxel of the (x, y, λ) datacube is viewed through a restricted set of angles, corresponding to the limited number of projected images captured by the camera. Due to its limited detector area and low diffraction efficiency at high diffraction orders, CTIS suffers from two missing cones in the spatio-spectral frequency domain. This limitation makes it challenging to image objects with flat spatial features and sharp spectral transitions.

Similarly, the tomographic approach has been demonstrated in ultrafast optical imaging to acquire an (x, y, t) datacube. Lai *et al.* developed a compressed ultrafast tomographic imaging (CUTI) approach utilizing a streak camera to shear an input (x, y, t) datacube along the time axis [50]. By varying the sweeping velocity of the streak camera, they captured multiple 2D projections of an (x, y, t) datacube and demonstrated tomographic reconstruction of an ultrafast scene at 0.5 trillion frames per second. However, unlike CTIS, which acquires all projections in parallel, CUTI performs measurements sequentially, capturing only one 2D projection at a time and thus requiring multiple shots to complete a measurement. Additionally, since sweeping can only be performed along a fixed spatial axis inside a streak camera, the sampling of an (x, y, t) datacube in the spatiotemporal frequency domain is further restricted, limiting the spatiotemporal features that can be recovered.

Dimension encoding: instead of capturing projections of a plenoptic function from multiple "views", the target dimension can be encoded using known signatures detectable by existing optical sensors. For instance, point spread function (PSF) engineering [83,84] can produce depth-variant PSFs and thus localize 3D fluorescent signals with a 2D widefield image. Similarly, relative depth can be inferred from changes in the polarization state of reflected light using a polarization camera [85]. Notably, coded aperture snapshot imaging represents a well-regarded approach in this domain. It uses a patterned mask to encode an image in the spatial domain. This is followed by shearing the high-dimensional light information along one spatial axis and compressively mapping it to a low-dimensional space in a superimposed manner. Similar to tomographic approaches, it has been successfully applied in light field [86–88], spectral [89–91], and ultrafast imaging [52,92–95].

Coded aperture snapshot spectral imaging (CASSI) is a representative hyperspectral imaging technique. CASSI encodes the input image with a random binary pattern using an absorption mask, then disperses the encoded image with a prism [51,96–99]. The spatio-spectrally multiplexed image is captured by a 2D camera. Image reconstruction involves solving the inverse problem of the image formation process. By employing algorithms such as

gradient projection for sparse reconstruction or a two-step iterative shrinkage/thresholding algorithm, Wagadarikar *et al.* [98] demonstrated that the spectral datacube can be reconstructed from such measurements. Operating on the principles of compressed sensing, CASSI requires the input scene to be sparse in the spatial gradient domain to function effectively.

Compressed ultrafast photography (CUP) [52] is the counterpart of CASSI in ultrafast imaging. By utilizing spatial encoding, CUP transforms a conventional 1D streak camera into a 2D snapshot ultrafast imaging device. The CUP system first captures an image of the object through a camera lens and relays this intermediate image to a spatial encoding device, typically a digital micromirror device (DMD), displaying a pseudo-random pattern. Light reflected from the "on" micromirrors is collected by a microscope objective and reimaged onto the entrance slit of a streak camera, which is fully opened to allow the formation of a 2D image on the photocathode. Inside the streak camera, this image is temporally sheared along the vertical axis by a varying voltage, with the amount of shearing determined by the time of arrival of the incident photons at a given voltage ramp rate. The final image is captured by a CCD in a single exposure. The image reconstruction process in CUP is similar to that of CASSI, allowing for a reasonable estimation of a time-lapse scene if the object is sparse in the spatial gradient domain. CUP has demonstrated versatility across various scientific fields, including fluorescence lifetime imaging microscopy [100], visualization of stochastic physical events [101–104], and imaging through scattering media.

B. Reducing Digital Layer Burden with Physical Components

1. In-Sensor Computation

In conventional optical imaging systems, there is a clear division between the physical sensor and digital processing components, necessitating substantial data transfer bandwidth and resulting in significant latency. To address this challenge, researchers have focused on integrating computational functionalities directly within the pixels of image sensors. This innovative approach allows for real-time processing and analysis of raw image data at the point of capture, thereby reducing the burden on external computational units (Fig. 5).

For example, Dudek *et al.* [105] designed the SCAMP-5 vision chip, which features a 256 × 256 processor array, with each pixel containing a programmable processing element. Unlike conventional image sensors that output raw images, the SCAMP-5 outputs the results of on-sensor computations, such as feature maps and optic flow maps. Building on this, So *et al.* [111] developed MantissaCam based on the SCAMP-5 vision chip. They implemented a perceptually inspired "mantissa" encoding scheme in each pixel, which compresses a high-dynamic-range (HDR) pixel value into a low-dynamic-range (LDR) one, with the HDR image subsequently recovered using a neural network. The SCAMP-5 sensor has also been programmed to achieve spatially varying pixel exposures [112], in-pixel recurrent neural networks [113], compressive light-field imaging [114], and hyperspectral imaging [114].

Beyond conventional CMOS technology, research is exploring new materials to enhance image sensor capabilities. Yang *et al*. [106] introduced in-sensor dynamic computing using multiple-terminal mixed-dimensional graphene–germanium (graphene–Ge) heterostructure device arrays to perform dynamic computing at the sensor level for intelligent machine vision. This sensor accurately extracts edge features of dim targets in various visual scenarios and robustly tracks these targets. Zhou *et al.* [115] developed a fully hardware-implemented sensor with versatile image processing functions, including contrast enhancement, background denoising, feature extraction, and high-level image recognition. This sensor utilizes emerging multimodal modified silk fibroin protein (MSFP)-based resistive memory arrays, simulating retinal cells for image pre-processing and the visual cortex for high-level image processing.

In addition to conventional electronic components, advanced optical components can also be integrated directly onto sensors, effectively reducing the volume of imaging systems and significantly enhancing their imaging capabilities. For instance, Zheng *et al.* [107] designed a multichannel meta-imager that collaborates with a digital backend to offload computationally expensive convolution operations into high-speed, low-power optics. Chen *et al.* [108] proposed an all-analog imaging chip combining electronic and light computing (ACCEL), which remarkably bypasses the need for analog-to-digital converters, achieving a computing latency of just 72 ns per frame.

Finally, in-pixel/sensor computing also revolutionizes image representations. Event cameras, or dynamic vision sensors, respond to local changes in brightness [116]. Recently, Huang *et al.* [109,110] developed a spike camera that is 1000 times faster

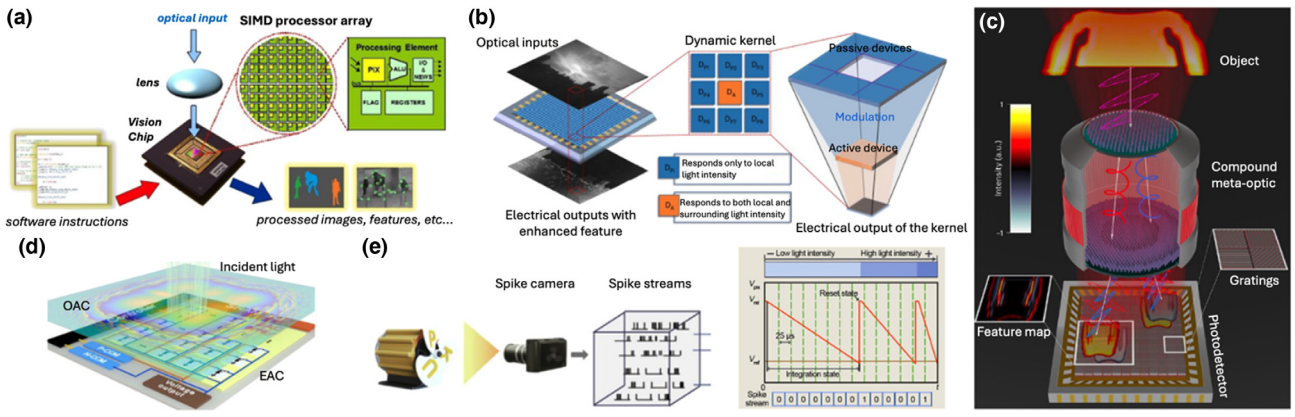

Fig. 5. In-sensor computing advancements for imaging. (a) SCAMP-5 vision chip [105]. (b) In-sensor dynamic computing [106]. (c) Multichannel meta-imager [107]. (d) ACCEL chip [108]. (e) Spike camera [109,110]. Figures are reproduced with permission from Ref. [105] The American Association for the Advancement of Science, Refs. [106–108] Springer Nature, and Ref. [109] the authors.

than conventional cameras using only consumer-level complementary metal–oxide–semiconductor (CMOS) sensors and integrated circuits. This camera represents vision data using a bit sequence array, where each bit indicates whether the accumulation of photons has reached a threshold.

2. Diffractive Optical Computation

Conventional digital processing units, limited by electronic bandwidth, often suffer from latency issues and consume substantial power to execute inference tasks. In contrast, diffractive optical computation leverages physical components, specifically diffractive optical elements (DOEs) or metasurfaces, to perform complex computational tasks directly within the optical domain. This approach significantly reduces the computational burden on digital processing units.

DOEs and metasurfaces manipulate the phase and amplitude of incoming light to perform specific computational tasks, such as image recognition, edge detection, and feature extraction. These components are designed to transform the light wavefronts so that the output is not a conventional image but rather a high-level representation of the object, such as a feature map or a processed data array. One of the significant advantages of diffractive optical computation is its ability to execute these tasks almost instantaneously as light travels through the optical elements. This eliminates the need for extensive data transfer and processing by digital units, leading to substantial power savings and faster computation.

Although diffractive optical computation has long been demonstrated to perform simple linear transformation tasks like Fourier transformation, recent breakthroughs have highlighted its potential in deep learning inference (Fig. 6).

A notable approach is the development of diffractive deep neural networks (D2NNs) [122], which can execute various machine learning and computer vision tasks traditionally managed by graphics processing units. D2NN handles analog optical waves propagating through a series of engineered surfaces like DOEs or metasurfaces to perform statistical inference. The transmittance coefficients of the diffractive units within the free-space optical processor are optimized for the specific machine learning task and the detector configuration at the output plane. After the training phase, the designed diffractive surfaces are fabricated and assembled to form a physical diffractive optical processor.

The early experimental demonstrations of D2NN were performed using THz wavelengths for simple tasks like object detection and target classification. Since these initial proof-of-concept demonstrations, significant design advances have been made, enhancing D2NN's capabilities to perform more complex functions, including imaging through unknown random diffusers [117,123], selectively imaging specific targets while blocking others [118], and hologram reconstruction [124]. The wavelength range has also been extended to visible light through the use of metasurfaces [119]. Furthermore, by tailoring the diffractive patterns, multiple computational tasks can be encoded into a same diffractive network, further improving the system's efficiency and capability [125,126]. For example, Li *et al.* [125] reported a polarization-multiplexed diffractive optical network that can perform a group of arbitrary linear transformations using a common set of diffractive layers.

Despite gaining wide popularity, D2NN implemented in free space typically employs linear materials for computing. Therefore, it faces challenges in performing nonlinear activations like ReLU (rectified linear unit). To address this problem, researchers have been exploring various strategies, such as incorporating nonlinear optical materials or hybrid approaches that combine optical and electronic processing [127]. For instance, Wang *et al.* [120] reported a nonlinear, multilayer optical neural network (ONN) encoder for image sensing, utilizing a commercial image intensifier as an optical-to-optical nonlinear activation function. This innovative approach allowed the nonlinear ONN to outperform similarly sized linear optical encoders across several representative tasks. In another example, Xia *et al.* [121] leveraged multiple scattering within a reverberating cavity to passively induce optical nonlinear

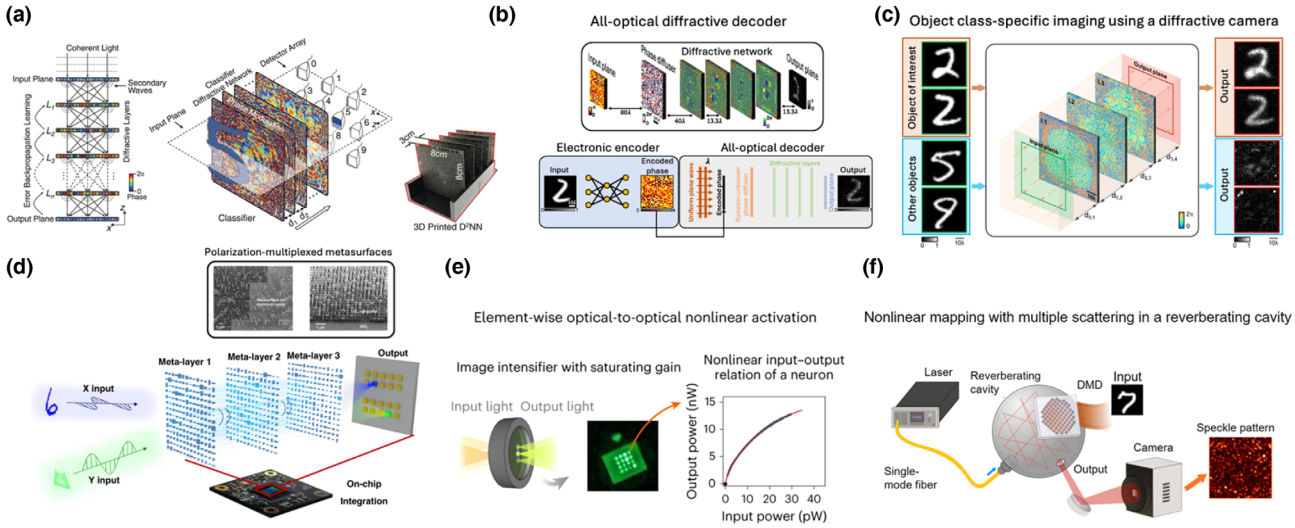

Fig. 6. Optical neural network for image processing tasks. (a) Diffractive deep neural network (D2NN) comprises multiple layers with complex-valued transmission coefficients [122]. (b) Hybrid electronic–optical encoder–decoder restores the information transmitted through random, unknown phase diffusers [117]. (c) Diffractive neural network performs object class-specific imaging with instantaneous all-optical erasure of other object classes [118]. (d) Polarization-multiplexed metasurface classifies objects using a CMOS camera chip in the visible spectrum [119]. (e) Optical nonlinear activation layer realized with a saturating image intensifier, resembling the sigmoid activation function [120]. (f) Nonlinear mapping between input (digital micromirror device) and output speckle pattern using an integrating sphere [121]. Figures are reproduced with permission from respective references.

random mapping, thereby eliminating the need for the high laser power typically required for activating nonlinear optical materials. By introducing nonlinearity to the optical domain, these methods enhanced the computational capabilities of D2NN, enabling it to perform a wider range of machine learning applications.

4. EXPLORING END-TO-END OPTIMIZATION-SYNERGY OF PHYSICAL AND DIGITAL DOMAINS

Instead of merely complementing each other's functions as discussed in Section 3, physical and digital layers can be co-optimized in an end-to-end manner. End-to-end joint optimization [128] aims to design optics with a merit function directly derived from the given visual task. This is achieved by creating a differentiable model of the entire imaging system and computational algorithms. The output of the algorithms is compared to the ground truth, and the error is backpropagated to adjust the parameters of both the hardware and the algorithm. The hardware design is then fabricated to perform in the real world with optimized capability for the targeted task.

The fusion of customized hardware and state-of-the-art algorithms, underscored by end-to-end joint optimization and principles of compressive sensing, stands to drastically reduce hardware requirements, streamline computations, and elevate the precision of imaging [129]. Conceptually, customized physical optics allow for the modulation of optical waves to encode implicit scene information into PSFs, which can then be decoded using reconstruction algorithms (Fig. 7).

A. Overview of Differentiable Optics Modeling

Although extensively studied for centuries, typical camera optical systems and image processing algorithms have been designed separately and sequentially. Existing approaches often rely on heuristic design or proxy metrics based on the PSF, rather than considering the final image quality after post-processing. Once the optics is fixed, the parameters of the image processing algorithms are tuned to achieve desirable image reproduction. Without a true end-to-end flow of joint optimization, it remains elusive to find an optimal computational camera for a given visual task. This joint design concept has long been at the core of computational optical imaging, and recent advancements in computational tools now enable efficient interpretation of a true end-to-end imaging process through machine learning techniques.

The resulting optics design paradigm, referred to as differentiable (deep) optics [128], stands out in the research community. Differentiable optics is an emerging field that combines principles from both physical optics and digital image processing, focusing on the development and application of differentiable models that seamlessly integrate physical optics, such as lenses, with computational algorithms, such as neural network models. By leveraging the synergy between these domains, researchers aim to enhance the capabilities of tailored optical systems, enable new imaging modalities, and improve the performance of various visual tasks subject to specific requirements. Fundamentally, it involves accurately modeling PSFs.

1. Wave Optics Representation

Light propagation through optical systems is commonly modeled using wave optics due to its effectiveness and simplicity [133,134]. In wave optics, light is described as an electromagnetic field in complex function form: $u_0(\xi, \eta) = A(\xi, \eta)\exp[j\phi(\xi, \eta)]$, where A and ϕ denote the amplitude and phase, respectively, that vary across space and $j = \sqrt{-1}$. For example, a plane wave propagating along the axial direction exhibits a uniform amplitude and phase across the (ξ, η) plane, whereas a spherical wave has an amplitude given by A/r and a phase given by $\phi = kr$, where $r = \sqrt{\xi^2 + \eta^2}$ and $k = 2\pi/\lambda$ is the wave number associated with the wavelength.

Spherical wavefronts are often employed to compute the PSF of imaging systems, crucial for analyzing system behavior. Wave optics also finds extensive applications in fields requiring high resolution and sensitivity such as microscopy, lithography, and holography [135–138]. Understanding the phase of light yields valuable insights into system performance and capabilities.

2. Wavefront Propagation

While the Rayleigh–Sommerfeld (RS) solution is a direct formulation of wavefront propagation, the more widely used formulation in scalar diffraction theory is the angular spectrum method (ASM) due to its simplicity in analyzing complex wavefronts in the Fourier domain [133]. Specifically, let the source field be $u_0(\xi, \eta)$; one can relate it to a target field $u_z(x, y)$ at a distance z and parallel to the source plane by a series of Fourier transforms, as follows:

$$\boldsymbol{u_z(x, y) = \mathcal{F}^{-1}\left[U(f_X, f_Y)\, H_z(f_X, f_Y)\right]},$$

where $U(f_X, f_Y) = \mathcal{F}[u_0(\xi, \eta)]$ is the angular spectrum of the source field, and the transfer function $H_z(f_X, f_Y) = \exp[jkz\sqrt{1 - (\lambda f_X)^2 - (\lambda f_Y)^2}]$ connects the source and target in the Fourier domain (f_X, f_Y). Here, $\mathcal{F}$ is the Fourier transform operator.

In certain scenarios, ASM can be reduced to simpler forms using approximations. One of the most frequently employed is the Fresnel approximation, which assumes near-field and near-axis propagation, linking the source and target via a single Fourier transform:

$$\boldsymbol{u_z(x, y) = \frac{e^{jkz}}{j\lambda z} e^{j\frac{k}{2z}(x^2+y^2)} \mathcal{F}\left[u_0(\xi, \eta)\, e^{j\frac{k}{2z}(\xi^2+\eta^2)}\right]}.$$

The near-field assumption holds true in most contemporary imaging systems, while the near-axis assumption applies when the system exhibits approximate shift-invariance [134,139,140]. In such cases, forward modeling of an imaging system simplifies to a straightforward convolution of the PSF with the imaging target I_{target}:

$$\boldsymbol{I_{detect}(x, y) = |u_z(x, y)|^2 * I_{target}(x, y)},$$

where $u_z(x, y)$ is derived by taking a spherical wave $u_0 = A/r\exp[jkr]$ into the equation, and $I_{\text{detect}}(x, y)$ denotes the raw input to the imaging sensor. This forward modeling allows us to manipulate the complex source field $u_0(\xi, \eta)$ to encode information about the scene into the sensor image tailored for domain-specific tasks.

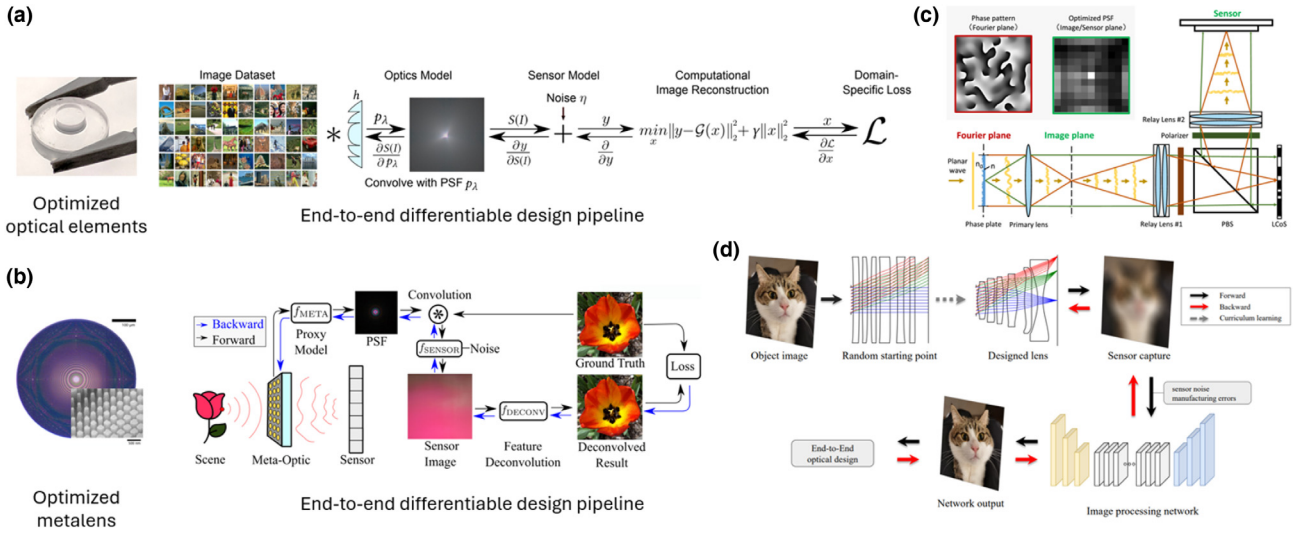

Fig. 7. End-to-end optimization of optics and image processing. (a) Differentiable imaging pipeline optimizes refractive/diffractive optical elements and image reconstruction algorithms simultaneously [128]. (b) Metasurface lens designed using an end-to-end framework [130]. (c) Snapshot compressed super-resolution imager with end-to-end optimized spatial (phase plate) and temporal (LCoS) encoding [131]. (d) End-to-end automated optical lens design [132]. Figures are reproduced with permission from respective references.

B. Overview of Differentiable Optics Modeling

The visual systems of animals have evolved to adapt to their specific environments. Interestingly, despite being utilized for a diverse range of applications, typical camera optics have primarily been engineered to mimic the human eye [128]. This raises an important, unanswered question: what is the optimal camera (optics) design for a given visual task? Consequently, application-domain-specific cameras that combine customized optics with modern image recovery algorithms are of rapidly growing interest in research. These camera designs hold immense potential and find applications in various fields, including ultrathin cameras for the internet-of-things (IoTs), robots, or drones, as well as computational cameras for microscopy and scientific imaging.

By designing a differentiable wave optics simulator mapping true source images to reconstructed ones, the encoder can be parameterized by the PSFs of physical optics, while the decoder, i.e., the image processing unit, can be a deep neural network. In this context, the parameters of this network model are jointly updated with the optics parameters.

Constructing the most proper loss functions and incorporating practical constraints are essential for optimizing optics. However, conventional design paradigms often restrict both optics and image processing units to being optimized only according to the quality of system modulation transfer functions (MTFs) or PSFs, failing to address the complex constraints required in various applications. The auto-differentiation-empowered wave propagation and optimization allows for tailoring loss functions flexibly for different applications, thereby practical fabrication and/or assembling metrics can be applied, including feature size, aspect ratio, and misalignment.

C. Representative Computational Imaging Tasks

The encoder–decoder deep optics imaging paradigm has received considerable attention in recent years, particularly in the context of snapshot computational imaging systems.

1. Extended Depth-of-Field Imaging

Extending depth of field has been a long-standing, representative application of computational imaging, as all-in-focus images are crucial in many scenarios, ranging from medical diagnostics to security monitoring. Early methods focused on incorporating hand-crafted phase masks to modify incoherent optical systems [141]. Later on, combining diffractive and refractive optical elements was proposed as a compromise to minimize chromatic aberrations in EDOF [142,143].

In recent years, the end-to-end (E2E) optimization paradigm has significantly enhanced the capabilities of EDOF techniques, treating the sensor as a bottleneck that integrates the incident wave quantities [144]. The concept of deep optics was first comprehensively explored in 2018 to realize achromatic EDOF and super-resolution imaging, demonstrating state-of-the-art performance using fabricated DOEs [128]. This breakthrough study highlighted the effectiveness of co-designing camera optics and reconstruction neural networks using a fully differentiable image formation model. However, the final image quality did not match the level achievable with off-the-shelf camera optics, mainly due to limitations associated with involving only a single optical surface. Subsequently, Akpinar *et al.* [145] investigated an optical model consisting of a refractive lens and a phase-modulating element built upon a DOE. Further, Sun *et al*. [146] developed a differentiable complex model and accordingly an end-to-end framework that could offer greater design flexibility over compact lens designs.

It is worth noting that there still exists a gap between theoretical solutions and real-life physical implementations, arising from various factors, including artifacts in manufacturing DOEs and limitations in the accuracy of the digital camera model and calibration process [130]. To bridge this gap, considerable efforts have been made in the field of hardware encoding designs [147,148], although the details of these efforts are beyond the scope of this discussion.

2. Compressive Hyperspectral Imaging

Lightweight DOE-empowered compressive hyperspectral imaging approaches have been extensively explored [149–152]. In particular, Peng *et al*. [152] proposed integrating optimized DOEs with post-capture processing algorithms, allowing for crafting lightweight optics capable of capturing the full visible spectrum. Although its design flow is not fully end-to-end, this approach opens possibilities for co-designing optics and computational algorithms in spectral imaging. Later on, given the insight that compact hyperspectral imaging can be realized by facilitating a spectrally rotating PSF from a customized DOE placed in front of the sensor, the deep optics paradigm has been introduced for designing a number of diffractive snapshot hyperspectral imaging systems [140,149].

In a recent work by Arguello *et al*. [153], a generalized deep optics approach has been proposed, involving the joint optimization of the DOE, color-coded apertures (CCAs), and decoder neural networks. Both the DOE and CCAs are designed to enhance freedom in spectral modulation. Notably, the exploration of deep optics frameworks for snapshot hyperspectral imaging is ongoing, with researchers investigating more generalized applications such as object detection in the short-wave infrared band [154].

3. Alternative Photography and Scientific Imaging

The E2E optimization paradigm has also made significant contributions to various areas of photography and scientific imaging, including high-dynamic-range (HDR) imaging [112,155], microscope imaging [156,157], large field-of-view (FoV) simple lens imaging [158], depth estimation [159,160], compressed imaging [131], and single-photon avalanche diode (SPAD) cameras [161]. These works aim to connect optics with downstream algorithms to achieve improved or novel imaging capabilities. For instance, Sun *et al.* [161] proposed an optically coded super-resolution technique with deep optics for SPAD cameras, achieving several orders higher resolution than practical single-photon imagers with exposure times of less than 1 ns. Shi *et al*. [162] also proposed to recover unobstructed scenes using a thin lens and a learnable DOE, enabling imaging through occlusions without inpainting.

It is worth noting that, in addition to conventional DOEs, various hardware encoders including refractive optics [132], microlens masks [159], and neural sensors [112] have been utilized for E2E designing domain-specific imagers, providing more freedom and feasibility in hardware parameters. For instance, unlike passive imaging models, Mask-ToF [159] uses a differentiable ToF simulator to jointly learn an optimal mask pattern and a refinement network for specialized depth map acquisition. In the field of video compressive imaging, Wang *et al*. [163] investigate the learnable structural mask, sensor response, and reconstruction algorithm. As a result, the incompatibility between dynamic range and temporal multiplexing can be greatly mitigated.

Narrowing the gap between hardware simulation design and actual optics fabrication is a crucial yet challenging issue [164]. For instance, conventional etching-based methods [165] are commonly leveraged for DOE prototyping, but they suffer from surface roughness and alignment issues between layers, which may cause various etching artifacts [166]. To address these challenges, several DOE design and manufacturing attempts have been conducted [167–169], yet there remains a certain distance to be traversed in the near term. It is expected that more remarkable outcomes will be seen in a wider range of innovative 2D or 3D imaging tasks as manufacturing procedures for customized optical/sensor components continue to advance.

D. Off-Axis Diffraction Modeling Further Bridges the Gap

Until now, discussions on computational optics design have primarily assumed shift-invariance, despite this assumption not always holding true in optical systems. In practice, higher-level aberrations manifest at larger diffraction angles, and the accuracy of the Fresnel propagation approximation in the previous equations decreases. As a result, shift-variant PSF modeling is necessary, requiring the use of the ASM. However, simulating the ASM in such cases demands significant computational resources. Recent advances have accelerated the implementation of shift-variant PSF modeling, from zero-padded fast Fourier transform (FFT) to the band-extended method [133,170]. The latter rearranges the sampling points to the effective band, a particularly noteworthy approach since the reason for introducing zero-paddings in the Fourier domain is often overlooked.

A recent advancement is the least-sampling angular spectrum method (LS-ASM) [171], which offers efficient and differentiable off-axis diffraction modeling with high accuracy. By utilizing the Fourier transform's shifting property, this method transforms off-axis diffraction into a quasi-on-axis condition and establishes minimum necessary sampling criteria for ASM at all angles. Specifically, the LS-ASM introduces a compensation factor to the source's oblique field, aligning the angular spectrum around zero frequency, expressed as

$$\boldsymbol{u}_{\textbf{on-axis}}(\boldsymbol{\xi}, \boldsymbol{\eta}) = \boldsymbol{u}_{\boldsymbol{0}}(\boldsymbol{\xi}, \boldsymbol{\eta}) - \boldsymbol{2\pi}\left(\boldsymbol{f}_{\boldsymbol{X}}^{\boldsymbol{c}}\boldsymbol{\xi} + \boldsymbol{f}_{\boldsymbol{Y}}^{\boldsymbol{c}}\boldsymbol{\eta}\right),$$

where f_X^c and f_Y^c are the original center of the angular spectrum. The final target field is obtained by propagating $u_{\mathrm{on-axis}}(\xi, \eta)$ along the axis. By removing the frequency offset, the minimum sampling rate is determined primarily by the effective bandwidth of the field, which establishes minimal sampling bounds for both spatial and frequency domains.

This method, demonstrated with an exemplary coded aperture imaging system [171], achieves a substantial speed boost, around 36× over the state-of-the-art [170] at 20°, allowing for high-frequency off-axis wavefront propagation to be computed within seconds on a commercial computer. Overall, LS-ASM presents a viable approach for integrating off-axis modeling into comprehensive deep optics applications.

5. GRAND CHALLENGES AND PERSPECTIVES

A. Accountable AI for Scientific Imaging

When utilizing digital layers to complement physical measurements, such as super-resolution or denoising, a common strategy is to employ data-driven approaches. These methods involve training a neural network to map the input measurement data directly to the object space, eliminating the need for a forward image model. However, these deep learning methods function as black boxes and are often constrained by the quality and diversity of their training datasets. There are limited guarantees regarding when a deep network will accurately resolve an imaging inverse problem versus when it might hallucinate information. If the training dataset is

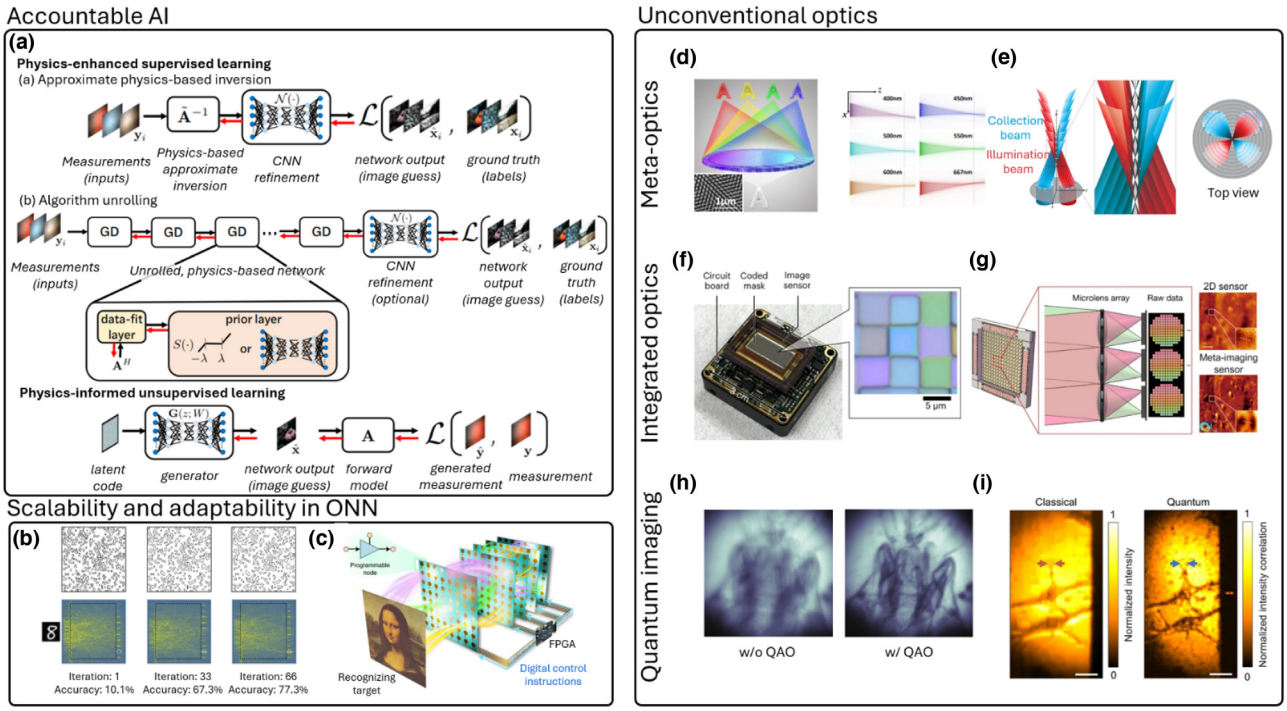

Fig. 8. Grand challenges and perspectives. (a) Physics-informed AI embeds physical principles and constraints into neural network training [174]. (b) Nanophotonic neural medium performs image recognition post-training [176]. (c) Array of programmable metasurfaces controlled by FPGAs makes a real-time and re-trainable intelligent machine [177]. (d) Ultracompact spectral light-field imaging [178]. (e) Bijective illumination collection imaging [179]. (f) Hyperspectral imaging with Fabry–Pérot filters directly placed onto the surface of a monochromatic image sensor [180]. (g) Aberration-free meta-imaging sensor integrating microlens array and piezo stage onto a conventional CMOS sensor [181]. (h) Quantum-assisted adaptive optical imaging (QAO) [182]. (i) Quantum microscopy [183]. Figures are reproduced with permission from respective references.

too restrictive, the network may fail to generalize effectively to other types of data, potentially compromising the reliability of the results. Therefore, ensuring accountability in AI models for scientific imaging remains an open challenge in the computational imaging field.

One effective remedy is to incorporate physics-based knowledge into the neural network architecture. By embedding physical principles and constraints into the network, it alleviates the burden on the neural network to learn these aspects from scratch. This approach, often referred to as physics-informed neural networks (PINNs) or hybrid models, allows the network to leverage established scientific knowledge, improving its learning efficiency and accuracy [172–175].

In PINN-based computational imaging [Fig. 8(a)], physical principles are typically incorporated by embedding the forward image model into a data fidelity term during image reconstruction. This allows the network to concentrate on learning the complex, nuanced aspects of the data that traditional data-driven models struggle to capture. Consequently, the network's predictions turn to be more reliable and interpretable, as they are anchored in well-understood physical principles. Moreover, this hybrid approach enhances the network's ability to generalize across various types of data, reducing the risk of overfitting to the training dataset. For example, unrolling classical iterative algorithms with neural networks and imposing measurement constraints from a known sensing matrix at each step promise interpretability and generalizability in the reconstruction of snapshot compressive imaging [184,185] and lensless imaging [186]. Wang *et al*. also demonstrated that incorporating a real-world diffraction model into a deep neural network can reconstruct phase images without labeled data for training [187].

Another strategy to enhance the accountability of AI in scientific imaging is to provide an uncertainty estimate during inference [188]. This approach involves quantifying the confidence in the neural network's output, offering a means to evaluate the reliability of the results quantitatively. In computational imaging, uncertainty quantification can provide confidence bounds for machine-learning-based image reconstruction tasks, thereby adding credibility to the predictions.

To this end, several techniques can be employed to estimate uncertainty in neural network models, such as Bayesian neural networks, Monte Carlo dropout, and ensemble methods [189]. Bayesian neural networks incorporate probability distributions into the model parameters, allowing for the calculation of uncertainty directly from the posterior distributions. Monte Carlo dropout, on the other hand, involves performing multiple forward passes with dropout enabled during inference, providing a measure of the variability in the predictions. Ensemble methods combine the predictions of multiple independently trained models to estimate uncertainty, with the variance among the models' outputs serving as an indicator of uncertainty.

Moreover, incorporating uncertainty estimates can aid in identifying regions of the data where the model is less certain, guiding further investigation or extra data collection. It can also highlight potential areas where the model may be prone to errors, enabling researchers to address these weaknesses proactively. For example, in a recent study, Ye *et al.* [190] incorporated uncertainty quantification into an optimized denoising model to guide adaptive

multiphoton microscopy image acquisition. They demonstrated that their method could maintain fine features in the denoised image while outperforming other denoising methods. This was achieved by adaptively selecting the most uncertain pixels for reimaging in a human endometrium tissue sample.

B. Scalability and Adaptability of Physical Domain Computing

When using physical components to partially or entirely replace digital layers, scalability often presents a significant challenge. For instance, implementing a DNN with an in-sensor computing paradigm necessitates a large number of interconnecting nodes. This requirement significantly increases the sensor complexity, posing substantial challenges for interconnect routing and layout design.

For free-space optical computing, such as D2NN, scalability is primarily constrained by the space-bandwidth product (SBP) of the diffractive layers. The size of the diffractive features determines the number of neurons that can be integrated into a neural network layer. For example, the diffractive features used in free-space optical processors must have dimensions of approximately half the optical wavelength, which is greater than 100 nm even for visible light. This size is significantly larger than the state-of-the-art transistors in chips manufactured using the 3-nm process. Consequently, integrating a large number of diffractive features of this scale would inevitably result in a much larger optical processor. Moreover, unlike CPUs and GPUs that can be assembled into clusters to tackle advanced computational problems, a versatile strategy for assembling diffractive networks into large-scale clusters of free-space optical processors is still lacking. Such a strategy is highly desired to fully leverage the parallelism advantages of optical computing and build low-loss, large-scale neural networks.

One promising solution to achieve a high density of processing units in optical computing is to utilize light reflection in a continuous medium with a heterogeneous distribution of refractive indices. A representative approach within this category is the nanophotonic neural medium (NNM) [Fig. 8(b)] [176]. Traditional layer-based D2NNs focus on forward light propagation and typically neglect the back-reflected light from each diffractive layer. In contrast, NNMs model light propagation in a host medium, such as SiO_2, containing numerous inclusions like air holes or other materials with different refractive indices. These inclusions scatter light in both forward and backward directions, causing the input light beams to undergo multiple scattering events and mix with each other. This process resembles linear matrix multiplication in a digital neural network, where the locations and shapes of inclusions function as weight parameters. Additionally, nonlinear activation can be achieved using inclusions made of dye semiconductors or graphene saturable absorbers. The multiple scattering events significantly increase the interaction between light beams and individual scatterers (analogous to neurons in a neural network), enabling the "deep" processing of input optical information. This enhanced interaction allows the optical network to handle more complex machine learning tasks.

Moreover, this approach can be extended to 3D space by utilizing a diffractive volume, where each discrete voxel is spatially engineered to collaboratively process the input optical information. The 3D nature of the diffractive volume provides a more scalable platform for integrating many processing units, further enhancing the information processing capabilities of NNMs. However, approaches like NNMs need to model complex light propagation in a scattering medium, making the calculation of light interactions with scatterers and the optimization of their refractive indices during training computationally intensive and challenging with current simulation methods. Even if an optimal NNM can be computed, fabricating nanoscale structures in 3D space remains a formidable task.

In addition to scalability, the ability to adapt physical-domain computing systems for specific tasks poses a significant challenge. Unlike conventional CPUs or GPUs, which can be easily reprogrammed for various machine learning tasks, physical-domain computing typically relies on fixed optical architectures. Once the hardware is constructed, it is difficult to reconfigure it for a different task.

For example, as discussed in Section 3.A.3, an effective strategy to address the dimensionality mismatch problem is to utilize spatial encoding. Approaches in this category often use fixed printed absorption masks to encode images. Although the mask pattern can be optimized for a specific type of object using end-to-end optimization, it cannot be reused for other types of objects. Similarly, in D2NN, the diffractive layers are typically fabricated using 3D printing. These layers are tailored for a specific machine learning task and cannot be easily readapted for another task.

This lack of flexibility in reprogramming and adapting to new tasks limits the versatility of physical-domain computing systems. To address this problem, the hardware must have a certain degree of flexibility to allow parametric tuning. For instance, instead of using fixed absorption masks, one can employ programmable spatial light modulators, such as digital micro-mirror devices, to encode an image [163]. Similarly, in D2NN, instead of using fixed diffractive layers, one can utilize electrically field-programmable gate arrays (FPGAs), which have been recently demonstrated for GHz wave applications [Fig. 8(c)] [177]. These GHz wave processors can be reconfigured to perform a range of tasks, from matrix inversion to image classification and information encoding/decoding. This capability allows the optical processor to evolve and adapt to different tasks, achieving optimal performance in a dynamic environment. It is worth noting that while adding active light modulators to a physical computing system enhances its versatility to handle various tasks, it also leads to significantly higher power consumption compared to purely passive systems. This increase in power usage diminishes one of the major advantages of using optical processors for computation, which is their typically lower power consumption compared to electronic counterparts.

C. Unconventional-Optics Empowered Computational Imaging

The advancement of computational optical imaging is ultimately driven by the development of novel optical and photonics devices and advanced AI algorithms. Particularly, the emergence of free-form optics [191], meta-optics [192], and quantum imaging [193] shows great promise in revolutionizing this field, as they can overcome many limitations of conventional optical imaging hardware and enable new capabilities through integration with AI algorithms [Figs. 8(d)–8(i)].

Meta-optics utilizes metamaterials to control light in ways that are not possible with conventional optics. These materials can manipulate light at the subwavelength scale, enabling more precise manipulation of optical fields. They provide greater degrees of freedom in a more compact form compared to regular DOEs

[194,195]. This propels advances in meta-based imaging [130], holography [196], and relevant fields [197,198].

For instance, Hua *et al.* [178] demonstrate ultra-compact spectral light-field imaging using a transversely dispersive metalens array and a monochrome imaging sensor, achieving advanced imaging with 4-nm spectral resolution and near-diffraction-limit spatial resolution in a single snapshot. Zaidi *et al.* [199] showcase a compact Mueller matrix imaging system that acquires all 16 components of an object's spatially varying Mueller matrix in a single shot, utilizing metasurfaces for structured polarization illumination and analysis. Additionally, Pahlebaninezhad *et al.* [179] propose bijective illumination collection imaging, employing metasurfaces to decouple lateral resolution from the depth of focus, achieving tissue imaging at a wavelength of 1.3 μm with 3.2-μm lateral resolution maintained nearly intact over a 1.25-mm depth of focus. However, the substantial computational cost associated with the large design space poses significant challenges to simulations currently, limiting the applications of metasurfaces in large-scale imaging tasks.

Moreover, integrating advanced meta-optics components onto sensors can significantly reduce the volume of computational imaging systems and enhance their capabilities. For example, Cai *et al*. [200] demonstrate compact angle-resolved spectral imaging by packaging the photodetector and tunable metasurface into a miniaturized spectrometer with a footprint of $4 \times 4\ \mu m^2$, achieving a wavelength accuracy of 0.17 nm and spectral resolution of 0.4 nm. Yako *et al*. [180] integrate an array of CMOS-compatible Fabry–Pérot filters on a monochromatic image sensor to achieve per-pixel spatial–spectral encoding. This sensor, coupled with an AI-driven image reconstruction algorithm, achieves a remarkable frame rate of 34.4 fps while maintaining full high-definition resolution. Wu *et al*. [181] introduced an integrated scanning light-field imaging sensor (meta-imaging sensor) to achieve high-speed aberration-corrected three-dimensional photography for universal applications without additional hardware modifications. The meta-imaging sensor can correct multisite aberration across 1,000 arcseconds on an 80-cm ground-based telescope without reducing the acquisition speed. In another example, Yi *et al*. [201] present a 3D light-field sensor by integrating lithographically patterned perovskite nanocrystal arrays on a color CCD sensor, capable of directly measuring the specific angle of visible light or X-rays without microlens arrays. Finally, combined with freeform optics, meta-optics is poised to offer new compact functionalities [202].

On the other hand, quantum imaging harnesses the quantum properties of light and their interactions with the environment to surpass the limits of classical imaging. The recent advancements in quantum-inspired single-photon detection technologies, combined with a new wave of AI algorithms, have led to significant progress in low-light-flux imaging and sensing. One of the most representative works in this field is first-photon imaging [203]. In this approach, only the very first detected photon at each scan location is used for 3D imaging. By using the number of pulses until a photon is detected as an indirect measurement of reflectivity, along with a piecewise-smooth assumption for both reflectivity and depth, a 3D image of a scene can be faithfully produced with the aid of computation. This approach showcases an impressive synergy between computation and the quantum mechanism of single-photon detection in imaging 3D scenes under extremely low-flux illuminations.

Since this initial demonstration, quantum-empowered computational imaging techniques have been surging, opening new avenues in a variety of applications previously thought to be impossible. For example, Cameron *et al*. [182] propose quantum-assisted adaptive optical imaging with entangled photons (QAO), demonstrating the ability to remove aberrations in biological samples without a guide star, making it independent of the imaging modality and specimen under study. Zhang *et al*. [204] introduce quantum imaging by coincidence from entanglement (ICE), utilizing spatially and polarization-entangled photon pairs to offer higher signal-to-noise ratios, quantitative quantum birefringence imaging capability, and 25 times greater suppression of stray light compared to classical imaging. He *et al*. [183] present quantum microscopy by coincidence (QMC) with balanced pathlengths, achieving super-resolution imaging at the Heisenberg limit, with a two-fold resolution improvement and resistance to stray light up to 155 times stronger than classical signals.

In summary, the future trajectory of computational optical imaging realizations and approaches is anticipated to be defined by a persistent cycle of innovation, propelled by advancements in computational algorithms, hardware technology, and their cross-disciplinary partnerships. These developments hold the potential to empower intelligent imaging systems to sense and understand the world across various scales and dimensions—from facilitating the exploration of microscopic cellular processes to aiding in unraveling the mysteries of distant galaxies.

Funding. National Institutes of Health (R01HL165318, RF1NS128488, R35GM128761).

Acknowledgment. The authors would like to thank Laura Waller for constructive discussions.

Disclosures. The authors declare no conflicts of interest.

Data availability. No data were generated or analyzed in the presented research.

REFERENCES

1. R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 4th, global edition (Pearson Education, 2018).
2. Z. Zhang, "A flexible new technique for camera calibration," IEEE Trans. Pattern Anal. Mach. Intell. **22**, 1330–1334 (2000).
3. M. Makarkin and D. Bratashov, "State-of-the-art approaches for image deconvolution problems, including modern deep learning architectures," Micromachines **12**, 1558 (2021).
4. K. Monakhova, K. Yanny, N. Aggarwal, *et al.*, "Spectral DiffuserCam: lensless snapshot hyperspectral imaging with a spectral filter array," Optica **7**, 1298–1307 (2020).
5. B. Huang, H. Babcock, and X. Zhuang, "Breaking the diffraction barrier: super-resolution imaging of cells," Cell **143**, 1047–1058 (2010).
6. B. A. Wilt, J. E. Fitzgerald, and M. J. Schnitzer, "Photon shot noise limits on optical detection of neuronal spikes and estimation of spike timing," Biophys. J. **104**, 51–62 (2013).
7. A. K. LaViolette and C. Xu, "Shot noise limits on binary detection in multiphoton imaging," Biomed. Opt. Express **12**, 7033–7048 (2021).
8. L. Gao and L. V. Wang, "A review of snapshot multidimensional optical imaging: measuring photon tags in parallel," Phys. Rep. **616**, 1–37 (2016).
9. G. Wetzstein, I. Ihrke, D. Lanman, *et al.*, "Computational plenoptic imaging," in *ACM SIGGRAPH* (Association for Computing Machinery, 2012), pp. 1–265.
10. H. Mikami, L. Gao, and K. Goda, "Ultrafast optical imaging technology: principles and applications of emerging methods," Nanophotonics **5**, 497–509 (2016).
11. M. Lelek, M. T. Gyparaki, G. Beliu, *et al.*, "Single-molecule localization microscopy," Nat. Rev. Methods Primers **1**, 1–27 (2021).

12. G. Vicidomini, P. Bianchini, and A. Diaspro, "STED super-resolved microscopy," Nat. Methods **15**, 173–182 (2018).
13. D. C. Lepcha, B. Goyal, A. Dogra, *et al.*, "Image super-resolution: a comprehensive review, recent trends, challenges and applications," Inf. Fusion **91**, 230–260 (2023).
14. G. Zheng, C. Shen, S. Jiang, *et al.*, "Concept, implementations and applications of Fourier ptychography," Nat. Rev. Phys. **3**, 207–223 (2021).
15. G. Zheng, R. Horstmeyer, and C. Yang, "Wide-field, high-resolution Fourier ptychographic microscopy," Nat. Photonics **7**, 739–745 (2013).
16. F. Xu, Z. Wu, C. Tan, *et al.*, "Fourier ptychographic microscopy 10 years on: a review," Cells **13**, 324 (2024).
17. X. Chen, S. Zhong, Y. Hou, *et al.*, "Superresolution structured illumination microscopy reconstruction algorithms: a review," Light Sci. Appl. **12**, 172 (2023).
18. M. G. L. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy," J. Microsc. **198**, 82–87 (2000).
19. J. Sun, C. Zuo, L. Zhang, *et al.*, "Resolution-enhanced Fourier ptychographic microscopy based on high-numerical-aperture illuminations," Sci. Rep. **7**, 1187 (2017).
20. Z. F. Phillips, R. Eckert, and L. Waller, "Quasi-Dome: a self-calibrated high-NA LED illuminator for Fourier ptychography," in *Imaging and Applied Optics 2017 (3D, AIO, COSI, IS, MATH, pcAOP)* (Optica, 2017), p. IW4E.5.
21. O. Bunk, M. Dierolf, S. Kynde, *et al.*, "Influence of the overlap parameter on the convergence of the ptychographical iterative engine," Ultramicroscopy **108**, 481–487 (2008).
22. Y. Xiao, S. Wei, S. Xue, *et al.*, "High-speed Fourier ptychographic microscopy for quantitative phase imaging," Opt. Lett. **46**, 4785–4788 (2021).
23. X. Ou, R. Horstmeyer, C. Yang, *et al.*, "Quantitative phase imaging via Fourier ptychographic microscopy," Opt. Lett. **38**, 4845–4848 (2013).
24. Y. Shu, J. Sun, J. Lyu, *et al.*, "Adaptive optical quantitative phase imaging based on annular illumination Fourier ptychographic microscopy," PhotoniX **3**, 24 (2022).
25. L. Tian and L. Waller, "3D intensity and phase imaging from light field measurements in an LED array microscope," Optica **2**, 104–111 (2015).
26. R. Horstmeyer, J. Chung, X. Ou, *et al.*, "Diffraction tomography with Fourier ptychography," Optica **3**, 827–835 (2016).
27. U. S. Kamilov, I. N. Papadopoulos, M. H. Shoreh, *et al.*, "Learning approach to optical tomography," Optica **2**, 517–522 (2015).
28. S. Dong, R. Horstmeyer, R. Shiradkar, *et al.*, "Aperture-scanning Fourier ptychography for 3D refocusing and super-resolution macroscopic imaging," Opt. Express **22**, 13586–13599 (2014).
29. P. Song, S. Jiang, H. Zhang, *et al.*, "Super-resolution microscopy via ptychographic structured modulation of a diffuser," Opt. Lett. **44**, 3645–3648 (2019).
30. A. Matlock, A. Sentenac, P. C. Chaumet, *et al.*, "Inverse scattering for reflection intensity phase microscopy," Biomed. Opt. Express **11**, 911–926 (2020).
31. C. Yurdakul, O. Avci, A. Matlock, *et al.*, "High-throughput, high-resolution interferometric light microscopy of biological nanoparticles," ACS Nano **14**, 2002–2013 (2020).
32. C. Shen, A. C. S. Chan, J. Chung, *et al.*, "Computational aberration correction of VIS-NIR multispectral imaging microscopy based on Fourier ptychography," Opt. Express **27**, 24923–24937 (2019).
33. E. Mudry, K. Belkebir, J. Girard, *et al.*, "Structured illumination microscopy using unknown speckle patterns," Nat. Photonics **6**, 312–315 (2012).
34. M. G. L. Gustafsson, L. Shao, P. M. Carlton, *et al.*, "Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination," Biophys. J. **94**, 4957–4970 (2008).
35. M. G. L. Gustafsson, "Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution," Proc. Natl. Acad. Sci. USA **102**, 13081–13086 (2005).
36. L. Fan, F. Zhang, H. Fan, *et al.*, "Brief review of image denoising techniques," Vis. Comput. Ind. Biomed. Art. **2**, 7 (2019).
37. J. Chen, H. Sasaki, H. Lai, *et al.*, "Three-dimensional residual channel attention networks denoise and sharpen fluorescence microscopy image volumes," Nat. Methods **18**, 678–687 (2021).
38. M. Weigert, U. Schmidt, T. Boothe, *et al.*, "Content-aware image restoration: pushing the limits of fluorescence microscopy," Nat. Methods **15**, 1090–1097 (2018).
39. J. Lehtinen, J. Munkberg, J. Hasselgren, *et al.*, "Noise2Noise: learning image restoration without clean data," in *35th International Conference on Machine Learning (PMLR)* (2018), pp. 2965–2974.
40. A. Krull, T.-O. Buchholz, and F. Jug, "Noise2Void-learning denoising from single noisy images," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2019), pp. 2124–2132.
41. J. Batson and L. Royer, "Noise2Self: blind denoising by self-supervision," in *36th International Conference on Machine Learning (PMLR)* (2019), pp. 524–533.
42. X. Li, G. Zhang, J. Wu, *et al.*, "Reinforcing neuron extraction and spike inference in calcium imaging using deep self-supervised denoising," Nat. Methods **18**, 1395–1400 (2021).
43. X. Li, X. Hu, X. Chen, *et al.*, "Spatial redundancy transformer for self-supervised fluorescence image denoising," Nat. Comput. Sci. **3**, 1067–1080 (2023).
44. X. Li, Y. Li, Y. Zhou, *et al.*, "Real-time denoising enables high-sensitivity fluorescence time-lapse imaging beyond the shot-noise limit," Nat. Biotechnol. **41**, 282–292 (2022).
45. G. Zhang, X. Li, Y. Zhang, *et al.*, "Bio-friendly long-term subcellular dynamic recording by self-supervised image enhancement microscopy," Nat. Methods **20**, 1957–1970 (2023).
46. T. Pang, H. Zheng, Y. Quan, *et al.*, "Recorrupted-to-recorrupted: unsupervised deep learning for image denoising," *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2021), pp. 2043–2052.
47. C. Qiao, Y. Zeng, Q. Meng, *et al.*, "Zero-shot learning enables instant denoising and super-resolution in optical fluorescence microscopy," Nat. Commun. **15**, 4180 (2024).
48. X. Yuan, D. J. Brady, and A. K. Katsaggelos, "Snapshot compressive imaging: theory, algorithms, and applications," IEEE Signal Process Mag. **38**, 65–88 (2021).
49. N. Hagen and E. L. Dereniak, "Analysis of computed tomographic imaging spectrometers. I. Spatial and spectral resolution," Appl. Opt. **47**, F85–F95 (2008).
50. Y. Lai, R. Shang, C.-Y. Côté, *et al.*, "Compressed ultrafast tomographic imaging by passive spatiotemporal projections," Opt. Lett. **46**, 1788–1791 (2021).
51. Z. Meng, J. Ma, and X. Yuan, "End-to-end low cost compressive spectral imaging with spatial-spectral self-attention," *16th European Conference on Computer Vision (ECCV), Part XXIII*, Glasgow, UK, August 23–28, 2020 (Springer, 2020), pp. 187–204.
52. L. Gao, J. Liang, C. Li, *et al.*, "Single-shot compressed ultrafast photography at one hundred billion frames per second," Nature **516**, 74–77 (2014).
53. N. Viganò, H. D. Sarkissian, C. Herzog, *et al.*, "Tomographic approach for the quantitative scene reconstruction from light field images," Opt. Express **26**, 22574–22602 (2018).
54. N. Viganò, P. M. Gil, C. Herzog, *et al.*, "Advanced light-field refocusing through tomographic modeling of the photographed scene," Opt. Express **27**, 7834–7856 (2019).
55. M. Levoy, R. Ng, A. Adams, *et al.*, "Light field microscopy," ACM Trans. Graph. **25**, 924–934 (2006).
56. M. Broxton, L. Grosenick, S. Yang, *et al.*, "Wave optics theory and 3-D deconvolution for the light field microscope," Opt. Express **21**, 25418–25439 (2013).
57. R. Ng, M. Levoy, M. Brédif, *et al.*, "Light field photography with a hand-held plenoptic camera," (Stanford University, 2005).
58. A. Lumsdaine and T. Georgiev, "The focused plenoptic camera," in *IEEE International Conference on Computational Photography (ICCP)* (2009), pp. 1–8.
59. G. Scrofani, J. Sola-Pikabea, A. Llavador, *et al.*, "FIMic: design for ultimate 3D-integral microscopy of in-vivo biological samples," Biomed. Opt. Express **9**, 335–346 (2018).
60. X. Lin, J. Wu, G. Zheng, *et al.*, "Camera array based light field microscopy," Biomed. Opt. Express **6**, 3179–3189 (2015).
61. F. L. Liu, G. Kuo, N. Antipa, *et al.*, "Fourier DiffuserScope: single-shot 3D Fourier light field microscopy with a diffuser," Opt. Express **28**, 28969–28986 (2020).
62. Z. Cai, J. Chen, G. Pedrini, *et al.*, "Lensless light-field imaging through diffuser encoding," Light Sci. Appl. **9**, 143 (2020).

63. N. Antipa, S. Necula, R. Ng, *et al.*, "Single-shot diffuser-encoded light field imaging," in *IEEE International Conference on Computational Photography (ICCP)* (2016), pp. 1–11.
64. Z. Lu, J. Wu, H. Qiao, *et al.*, "Phase-space deconvolution for light field microscopy," Opt. Express **27**, 18131–18145 (2019).
65. S. Mahmoudpour, C. Pagliari, and P. Schelkens, "Learning-based light field imaging: an overview," EURASIP J. Image Video Process. **2024**, 12 (2024).
66. Z. Wang, L. Zhu, H. Zhang, *et al.*, "Real-time volumetric reconstruction of biological dynamics with light-field microscopy and deep learning," Nat. Methods **18**, 551–556 (2021).
67. C. Yi, Y. Ma, M. Sun, *et al.*, "High-fidelity, generalizable light-field reconstruction of biological dynamics with physics-informed meta neural representation," (2024).
68. N. Wagner, F. Beuttenmueller, N. Norlin, *et al.*, "Deep learning-enhanced light-field imaging with continuous validation," Nat. Methods **18**, 557–563 (2021).
69. Z. Lu, Y. Liu, M. Jin, *et al.*, "Virtual-scanning light-field microscopy for robust snapshot high-resolution volumetric imaging," Nat. Methods **20**, 735–746 (2023).
70. X. Feng and L. Gao, "Ultrafast light field tomography for snapshot transient and non-line-of-sight imaging," Nat. Commun. **12**, 2179 (2021).
71. Z. Wang, T. K. Hsiai, and L. Gao, "Augmented light field tomography through parallel spectral encoding," Optica **10**, 62–65 (2023).
72. X. Feng, Y. Ma, and L. Gao, "Compact light field photography towards versatile three-dimensional vision," Nat. Commun. **13**, 3333 (2022).
73. T. Okamoto and I. Yamaguchi, "Simultaneous acquisition of spectral image information," Opt. Lett. **16**, 1277–1279 (1991).
74. T. V. Bulygin and G. N. Vishnyakov, "Spectrotomography: a new method of obtaining spectrograms of two-dimensional objects," Proc. SPIE **1843**, 315–322 (1992).
75. N. Hagen, E. L. Dereniak, and D. T. Sass, "Maximizing the resolution of a CTIS instrument," Proc. SPIE **6302**, 168–178 (2006).
76. N. Hagen and E. L. Dereniak, "New grating designs for a CTIS imaging spectrometer," Proc. SPIE **6565**, 216–224 (2007).
77. C. E. Volin, "Portable snapshot infrared imaging spectrometer," Ph.D. (The University of Arizona, 2000).
78. L. Yuan, Q. Song, H. Liu, *et al.*, "Super-resolution computed tomography imaging spectrometry," Photon. Res. **11**, 212–224 (2023).
79. B. K. Ford, M. R. Descour, and R. M. Lynch, "Large-image-format computed tomography imaging spectrometer for fluorescence microscopy," Opt. Express **9**, 444–453 (2001).
80. C. E. Volin, B. K. Ford, M. R. Descour, *et al.*, "High-speed spectral imager for imaging transient fluorescence phenomena," Appl. Opt. **37**, 8112–8119 (1998).
81. E. K. Hege, D. O'Connell, W. Johnson, *et al.*, "Hyperspectral imaging for astronomy and space surveillance," Proc. SPIE **5159**, 380–391 (2004).
82. W. R. Johnson, D. W. Wilson, W. Fink, *et al.*, "Snapshot hyperspectral imaging in ophthalmology," J. Biomed. Opt. **12**, 014036 (2007).
83. S. R. P. Pavani, M. A. Thompson, J. S. Biteen, *et al.*, "Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function," Proc. Natl. Acad. Sci. USA **106**, 2995–2999 (2009).
84. Y. Shechtman, S. J. Sahl, A. S. Backer, *et al.*, "Optimal point spread function design for 3D imaging," Phys. Rev. Lett. **113**, 133902 (2014).
85. X. Li, Z. Liu, Y. Cai, *et al.*, "Polarization 3D imaging technology: a review," Front. Phys. **11**, 1198457 (2023).
86. K. Marwah, G. Wetzstein, Y. Bando, *et al.*, "Compressive light field photography using overcomplete dictionaries and optimized projections," ACM Trans. Graph. **32**, 46 (2013).
87. A. Ashok and M. A. Neifeld, "Compressive light field imaging," Proc. SPIE **7690**, 221–232 (2010).
88. S. D. Babacan, R. Ansorge, M. Luessi, *et al.*, "Compressive light field sensing," IEEE Trans. Image Process. **21**, 4746–4757 (2012).
89. G. R. Arce, D. J. Brady, L. Carin, *et al.*, "Compressive coded aperture spectral imaging: an introduction," IEEE Signal Process Mag. **31**, 105–115 (2014).
90. X. Cao, T. Yue, X. Lin, *et al.*, "Computational snapshot multispectral cameras: toward dynamic capture of the spectral world," IEEE Signal Process Mag. **33**, 95–108 (2016).
91. X. Lin, Y. Liu, J. Wu, *et al.*, "Spatial-spectral encoded compressive hyperspectral imaging," ACM Trans. Graph. **36**, 1–11 (2014).
92. P. Llull, X. Liao, X. Yuan, *et al.*, "Coded aperture compressive temporal imaging," Opt. Express **21**, 10526–10545 (2013).
93. P. Wang, J. Liang, and L. V. Wang, "Single-shot ultrafast imaging attaining 70 trillion frames per second," Nat. Commun. **11**, 2091 (2020).
94. D. Qi, S. Zhang, C. Yang, *et al.*, "Single-shot compressed ultrafast photography: a review," Adv. Photon. **2**, 014003 (2020).
95. J. Park and L. Gao, "Continuously streaming compressed high-speed photography using time delay integration," Optica **8**, 1620–1623 (2021).
96. A. Wagadarikar, R. John, R. Willett, *et al.*, "Single disperser design for coded aperture snapshot spectral imaging," Appl. Opt. **47**, B44–B51 (2008).
97. A. A. Wagadarikar, N. P. Pitsianis, X. Sun, *et al.*, "Spectral image estimation for coded aperture snapshot spectral imagers," Proc. SPIE **7076**, 707602 (2008).
98. A. A. Wagadarikar, N. P. Pitsianis, X. Sun, *et al.*, "Video rate spectral imaging using a coded aperture snapshot spectral imager," Opt. Express **17**, 6368–6388 (2009).
99. R. Zhao, C. Yang, R. T. Smith, *et al.*, "Coded aperture snapshot spectral imaging fundus camera," Sci. Rep. **13**, 12007 (2023).
100. Y. Ma, Y. Lee, C. Best-Popescu, *et al.*, "High-speed compressed-sensing fluorescence lifetime imaging microscopy of live cells," Proc. Natl. Acad. Sci. USA **118**, e2004176118 (2021).
101. J. Liang, C. Ma, L. Zhu, *et al.*, "Single-shot real-time video recording of a photonic Mach cone induced by a scattered light pulse," Sci. Adv. **3**, e1601814 (2017).
102. Y. Zhang, B. Shen, T. Wu, *et al.*, "Ultrafast and hypersensitive phase imaging of propagating internodal current flows in myelinated axons and electromagnetic pulses in dielectrics," Nat. Commun. **13**, 5247 (2022).
103. L. Fan, X. Yan, H. Wang, *et al.*, "Real-time observation and control of optical chaos," Sci. Adv. **7**, eabc8448 (2021).
104. J. C. Jing, X. Wei, and L. V. Wang, "Spatio-temporal-spectral imaging of non-repeatable dissipative soliton dynamics," Nat. Commun. **11**, 2059 (2020).
105. P. Dudek, T. Richardson, L. Bose, *et al.*, "Sensor-level computer vision with pixel processor arrays for agile robots," Science Robotics **7**(67), abl7755 (2022).
106. Y. Yang, C. Pan, Y. Li, *et al.*, "In-sensor dynamic computing for intelligent machine vision," Nat. Electron. **7**, 225–233 (2024).
107. H. Zheng, Q. Liu, I. I. Kravchenko, *et al.*, "Multichannel meta-imagers for accelerating machine vision," Nat. Nanotechnol. **19**, 471–478 (2024).
108. Y. Chen, M. Nazhamaiti, H. Xu, *et al.*, "All-analog photoelectronic chip for high-speed vision tasks," Nature **623**, 48–57 (2023).
109. T. Huang, Y. Zheng, Z. Yu, *et al.*, "1000x faster camera and machine vision with ordinary devices," Engineering **25**, 110–119 (2023).
110. Y. Zheng, J. Zhang, R. Zhao, *et al.*, "SpikeCV: open a continuous computer vision era," arXiv (2024).
111. H. M. So, J. N. P. Martel, G. Wetzstein, *et al.*, "MantissaCam: learning snapshot high-dynamic-range imaging with perceptually-based in-pixel irradiance encoding," in *IEEE International Conference on Computational Photography (ICCP)* (2022), pp. 1–12.
112. J. N. P. Martel, L. K. Müller, S. J. Carey, *et al.*, "Neural sensors: learning pixel exposures for HDR imaging and video compressive sensing with programmable sensors," IEEE Trans. Pattern Anal. Mach. Intell. **42**, 1642–1653 (2020).
113. H. M. So, L. Bose, P. Dudek, *et al.*, "PixelRNN: in-pixel recurrent neural networks for end-to-end-optimized perception with neural sensors," arXiv (2023).
114. E. Vargas, J. N. P. Martel, G. Wetzstein, *et al.*, "Time-multiplexed coded aperture imaging: learned coded aperture and pixel exposures for compressive imaging systems," arXiv (2021).
115. G. Zhou, J. Li, Q. Song, *et al.*, "Full hardware implementation of neuromorphic visual system based on multimodal optoelectronic resistive memory arrays for versatile image processing," Nat. Commun. **14**, 8489 (2023).
116. P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128\times 128 120 dB 15 μs latency asynchronous temporal contrast vision sensor," IEEE J. Solid-State Circuits **43**, 566–576 (2008).
117. Y. Li, T. Gan, B. Bai, *et al.*, "Optical information transfer through random unknown diffusers using electronic encoding and diffractive decoding," Adv. Photon. **5**, 046009 (2023).
118. B. Bai, Y. Luo, T. Gan, *et al.*, "To image, or not to image: class-specific diffractive cameras with all-optical erasure of undesired objects," eLight **2**, 14 (2022).

119. X. Luo, Y. Hu, X. Ou, *et al.*, "Metasurface-enabled on-chip multiplexed diffractive neural networks in the visible," Light Sci. Appl. **11**, 158 (2022).
120. T. Wang, M. M. Sohoni, L. G. Wright, *et al.*, "Image sensing with multilayer nonlinear optical neural networks," Nat. Photonics **17**, 408–415 (2023).
121. F. Xia, K. Kim, Y. Eliezer, *et al.*, "Nonlinear optical encoding enabled by recurrent linear scattering," Nat. Photon. **18**, 1067–1075 (2024).
122. X. Lin, Y. Rivenson, N. T. Yardimci, *et al.*, "All-optical machine learning using diffractive deep neural networks," Science **361**, 1004–1008 (2018).
123. Y. Luo, Y. Zhao, J. Li, *et al.*, "Computational imaging without a computer: seeing through random diffusers at the speed of light," eLight **2**, 4 (2022).
124. M. S. Sakib Rahman and A. Ozcan, "Computer-free, all-optical reconstruction of holograms using diffractive networks," ACS Photonics **8**, 3375–3384 (2021).
125. J. Li, Y.-C. Hung, O. Kulce, *et al.*, "Polarization multiplexed diffractive computing: all-optical implementation of a group of linear transformations through a polarization-encoded diffractive network," Light Sci. Appl. **11**, 153 (2022).
126. J. Li, T. Gan, B. Bai, *et al.*, "Massively parallel universal linear transformations using a wavelength-multiplexed diffractive optical network," Adv. Photon. **5**, 016003 (2023).
127. T. Yan, J. Wu, T. Zhou, *et al.*, "Fourier-space diffractive deep neural network," Phys. Rev. Lett. **123**, 023901 (2019).
128. V. Sitzmann, S. Diamond, Y. Peng, *et al.*, "End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging," ACM Trans. Graph. **37**, 114 (2018).
129. S. H. Chan, "Computational image formation: simulators in the deep learning era," J. Imaging Sci. Technol. **67**, 1–17 (2023).
130. E. Tseng, S. Colburn, J. Whitehead, *et al.*, "Neural nano-optics for high-quality thin lens imaging," Nat. Commun. **12**, 6493 (2021).
131. B. Zhang, X. Yuan, C. Deng, *et al.*, "End-to-end snapshot compressed super-resolution imaging with deep optics," Optica **9**, 451–454 (2022).
132. X. Yang, Q. Fu, and W. Heidrich, "Curriculum learning for ab initio deep learned refractive optics," Nat. Commun. **15**, 6572 (2024).
133. J. W. Goodman, *Introduction to Fourier Optics* (Roberts and Company, 2005).
134. H. Ikoma, C. M. Nguyen, C. A. Metzler, *et al.*, "Depth from defocus with learned optics for imaging and occlusion-aware depth estimation," in *IEEE International Conference on Computational Photography (ICCP)* (2021), pp. 1–12.
135. C. Zheng, G. Zhao, and P. So, "Close the design-to-manufacturing gap in computational optics with a "Real2Sim" learned two-photon neural lithography simulator," in *SIGGRAPH Asia* (Association for Computing Machinery, 2023), pp. 1–9.
136. Y. Peng, X. Dun, Q. Sun, *et al.*, "Mix-and-match holography," ACM Trans. Graph. **36**, 191 (2017).
137. S. Choi, M. Gopakumar, Y. Peng, *et al.*, "Neural 3D holography: learning accurate wave propagation models for 3D holographic virtual and augmented reality displays," ACM Trans. Graph. **40**, 240 (2021).
138. E. Bostan, R. Heckel, M. Chen, *et al.*, "Deep phase decoder: self-calibrating phase microscopy with an untrained deep neural network," Optica **7**, 559–562 (2020).
139. X. Liu, L. Li, X. Liu, *et al.*, "Investigating deep optics model representation in affecting resolved all-in-focus image quality and depth estimation fidelity," Opt. Express **30**, 36973–36984 (2022).
140. X. Dun, H. Ikoma, G. Wetzstein, *et al.*, "Learned rotationally symmetric diffractive achromat for full-spectrum computational imaging," Optica **7**, 913–922 (2020).
141. E. R. Dowski and W. T. Cathey, "Extended depth of field through wavefront coding," Appl. Opt. **34**, 1859–1866 (1995).
142. A. Flores, M. R. Wang, and J. J. Yang, "Achromatic hybrid refractive-diffractive lens with extended depth of focus," Appl. Opt. **43**, 5618–5630 (2004).
143. Z. Liu, "Diffractive lens with extended depth of focus and its applications," (University of Miami, 2007).
144. G. Wetzstein, A. Ozcan, S. Gigan, *et al.*, "Inference in artificial intelligence with deep optics and photonics," Nature **588**, 39–47 (2020).
145. U. Akpinar, E. Sahin, M. Meem, *et al.*, "Learning wavefront coding for extended depth of field imaging," IEEE Trans. Image Process. **30**, 3307–3320 (2021).
146. Q. Sun, C. Wang, Q. Fu, *et al.*, "End-to-end complex lens design with differentiate ray tracing," ACM Trans. Graph. **40**, 71 (2021).
147. S. Pinilla, S. R. M. Rostami, I. Shevkunov, *et al.*, "Hybrid diffractive optics design via hardware-in-the-loop methodology for achromatic extended-depth-of-field imaging," Opt. Express **30**, 32633–32649 (2022).
148. S. R. M. Rostami, S. Pinilla, I. Shevkunov, *et al.*, "On design of hybrid diffractive optics for achromatic extended depth-of-field (EDoF) RGB imaging," Proc. SPIE **12136**, 160–175 (2022).
149. S.-H. Baek, I. Kim, D. Gutierrez, *et al.*, "Compact single-shot hyperspectral imaging using a prism," ACM Trans. Graph. **36**, 1–12 (2017).
150. F. Heide, Q. Fu, Y. Peng, *et al.*, "Encoded diffractive optics for full-spectrum computational imaging," Sci. Rep. **6**, 33543 (2016).
151. D. S. Jeon, S.-H. Baek, S. Yi, *et al.*, "Compact snapshot hyperspectral imaging with diffracted rotation," ACM Trans. Graph. **38**, 117 (2019).
152. Y. Peng, Q. Fu, F. Heide, *et al.*, "The diffractive achromat full spectrum computational imaging with diffractive optics," ACM Trans. Graph. **35**, 31 (2016).
153. H. Arguello, S. Pinilla, Y. Peng, *et al.*, "Shift-variant color-coded diffractive spectral imaging system," Optica **8**, 1424–1434 (2021).
154. N. Wei, Y. Sun, T. Jiang, *et al.*, "Direct object detection with snapshot multispectral compressed imaging in a short-wave infrared band," Opt. Lett. **49**, 1941–1944 (2024).
155. C. A. Metzler, H. Ikoma, Y. Peng, *et al.*, "Deep optics for single-shot high-dynamic-range imaging," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2020), pp. 1375–1385.
156. E. Nehme, D. Freedman, R. Gordon, *et al.*, "DeepSTORM3D: dense 3D localization microscopy and PSF design by deep learning," Nat. Methods **17**, 734–740 (2020).
157. H. Ikoma, T. Kudo, Y. Peng, *et al.*, "Deep learning multi-shot 3D localization microscopy using hybrid optical–electronic computing," Opt. Lett. **46**, 6023–6026 (2021).
158. Y. Peng, Q. Sun, X. Dun, *et al.*, "Learned large field-of-view imaging with thin-plate optics," ACM Trans. Graph. **38**, 219 (2019).
159. I. Chugunov, S.-H. Baek, Q. Fu, *et al.*, "Mask-ToF: learning microlens masks for flying pixel correction in time-of-flight imaging," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2021), pp. 9112–9122.
160. Y. Wu, V. Boominathan, H. Chen, *et al.*, "PhaseCam3D — learning phase masks for passive single view depth estimation," in *IEEE International Conference on Computational Photography (ICCP)* (2019), pp. 1–12.
161. Q. Sun, J. Zhang, X. Dun, *et al.*, "End-to-end learned, optically coded super-resolution SPAD camera," ACM Trans. Graph. **39**, 9 (2020).
162. Z. Shi, Y. Bahat, S.-H. Baek, *et al.*, "Seeing through obstructions with diffractive cloaking," ACM Trans. Graph. **41**, 37 (2022).
163. P. Wang, L. Wang, and X. Yuan, "Deep optics for video snapshot compressive imaging," in *IEEE/CVF International Conference on Computer Vision* (2023), pp. 10646–10656.
164. H. Amata, Q. Fu, and W. Heidrich, "Additive fabrication of SiO_2-based micro-optics with lag-free depth and reduced roughness," Opt. Express **31**, 41533–41545 (2023).
165. D. C. O'Shea, T. J. Suleski, A. D. Kathman, *et al.*, *Diffractive Optics: Design, Fabrication, and Test* (SPIE, 2003).
166. F. Laermer, S. Franssila, L. Sainiemi, *et al.*, "Chapter 16-Deep reactive ion etching," in *Handbook of Silicon Based MEMS Materials and Technologies, Micro and Nano Technologies*, 3rd ed., M. Tilli, M. Paulasto-Krockel, M. Petzold, H. Theuss, T. Motooka, and V. Lindroos, eds. (Elsevier, 2020), pp. 417–446.
167. Q. Fu, H. Amata, B. Gerard, *et al.*, "Additive lithographic fabrication of a tilt-Gaussian-vortex mask for focal plane wavefront sensing," Proc. SPIE **11889**, 162–170 (2021).
168. Q. Fu, H. Amata, and W. Heidrich, "Etch-free additive lithographic fabrication methods for reflective and transmissive micro-optics," Opt. Express **29**, 36886–36899 (2021).
169. Y. Zheng, Q. Fu, H. Amata, *et al.*, "Hexagonal diffractive optical elements," Opt. Express **31**, 43864–43876 (2023).
170. W. Zhang, H. Zhang, K. Matsushima, *et al.*, "Shifted band-extended angular spectrum method for off-axis diffraction calculation," Opt. Express **29**, 10089–10103 (2021).
171. H. Wei, X. Liu, X. Hao, *et al.*, "Modeling off-axis diffraction with the least-sampling angular spectrum method," Optica **10**, 959–962 (2023).

172. G. E. Karniadakis, I. G. Kevrekidis, L. Lu, *et al.*, "Physics-informed machine learning," Nat. Rev. Phys. **3**, 422–440 (2021).
173. A. Kadambi, C. de Melo, C.-J. Hsieh, *et al.*, "Incorporating physics into data-driven computer vision," Nat. Mach. Intell. **5**, 572–580 (2023).
174. K. Monakhova, "Physics-informed machine learning for computational imaging," (University of California, 2022).
175. C. Zuo, J. Qian, S. Feng, *et al.*, "Deep learning in optical metrology: a review," Light Sci. Appl. **11**, 39 (2022).
176. E. Khoram, A. Chen, D. Liu, *et al.*, "Nanophotonic media for artificial neural inference," Photon. Res. **7**, 823–827 (2019).
177. C. Liu, Q. Ma, Z. J. Luo, *et al.*, "A programmable diffractive deep neural network based on a digital-coding metasurface array," Nat. Electron. **5**, 113–122 (2022).
178. X. Hua, Y. Wang, S. Wang, *et al.*, "Ultra-compact snapshot spectral light-field imaging," Nat. Commun. **13**, 2732 (2022).
179. M. Pahlevaninezhad, Y.-W. Huang, M. Pahlevani, *et al.*, "Metasurface-based bijective illumination collection imaging provides high-resolution tomography in three dimensions," Nat. Photonics **16**, 203–211 (2022).
180. M. Yako, Y. Yamaoka, T. Kiyohara, *et al.*, "Video-rate hyperspectral camera based on a CMOS-compatible random array of Fabry–Pérot filters," Nat. Photonics **17**, 218–223 (2023).
181. J. Wu, Y. Guo, C. Deng, *et al.*, "An integrated imaging sensor for aberration-corrected 3D photography," Nature **612**, 62–71 (2022).
182. P. Cameron, B. Courme, C. Vernière, *et al.*, "Adaptive optical imaging with entangled photons," Science **383**, 1142–1148 (2024).
183. Z. He, Y. Zhang, X. Tong, *et al.*, "Quantum microscopy of cells at the Heisenberg limit," Nat. Commun. **14**, 2441 (2023).
184. Z. Meng, X. Yuan, and S. Jalali, "Deep unfolding for snapshot compressive imaging," Int. J. Comput. Vis. **131**, 2933–2958 (2023).
185. J. Zhang and B. Ghanem, "ISTA-Net: interpretable optimization-inspired deep network for image compressive sensing," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2018), pp. 1828–1837.
186. K. Monakhova, J. Yurtsever, G. Kuo, *et al.*, "Learned reconstructions for practical mask-based lensless imaging," Opt. Express **27**, 28075–28090 (2019).
187. F. Wang, Y. Bian, H. Wang, *et al.*, "Phase imaging with an untrained neural network," Light Sci. Appl. **9**, 77 (2020).
188. M. Abdar, F. Pourpanah, S. Hussain, *et al.*, "A review of uncertainty quantification in deep learning: techniques, applications and challenges," Inf. Fusion **76**, 243–297 (2021).
189. H. M. D. Kabir, A. Khosravi, M. A. Hosen, *et al.*, "Neural network-based uncertainty quantification: a survey of methodologies and applications," IEEE Access **6**, 36218–36234 (2018).
190. C. T. Ye, J. Han, K. Liu, *et al.*, "Learned, uncertainty-driven adaptive acquisition for photon-efficient multiphoton microscopy," arXiv (2023).
191. J. P. Rolland, M. A. Davies, T. J. Suleski, *et al.*, "Freeform optics for imaging," Optica **8**, 161–176 (2021).
192. K. Ou, H. Wan, G. Wang, *et al.*, "Advances in meta-optics and meta-surfaces: fundamentals and applications," Nanomaterials **13**, 1235 (2023).
193. M. Padgett, "Quantum imaging overview," Proc. SPIE **12447**, 1244702 (2023).
194. M. K. Chen, X. Liu, Y. Sun, *et al.*, "Artificial intelligence in meta-optics," Chem. Rev. **122**, 15356–15413 (2022).
195. S. Banerji, M. Meem, A. Majumder, *et al.*, "Imaging with flat optics: metalenses or diffractive lenses?" Optica **6**, 805–810 (2019).
196. G. Zheng, H. Mühlenbernd, M. Kenney, *et al.*, "Metasurface holograms reaching 80% efficiency," Nat. Nanotechnol. **10**, 308–312 (2015).
197. Z. Liu, D. Wang, H. Gao, *et al.*, "Metasurface-enabled augmented reality display: a review," Adv. Photon. **5**, 034001 (2023).
198. F. Zhao, R. Lu, X. Chen, *et al.*, "Metalens-assisted system for underwater imaging," Laser Photon. Rev. **15**, 2100097 (2021).
199. A. Zaidi, N. A. Rubin, M. L. Meretska, *et al.*, "Metasurface-enabled single-shot and complete Mueller matrix imaging," Nat. Photonics **18**, 704–712 (2024).
200. G. Cai, Y. Li, Y. Zhang, *et al.*, "Compact angle-resolved metasurface spectrometer," Nat. Mater. **23**, 71–78 (2024).
201. L. Yi, B. Hou, H. Zhao, *et al.*, "X-ray-to-visible light-field detection through pixelated colour conversion," Nature **618**, 281–286 (2023).
202. D. K. Nikolov, A. Bauer, F. Cheng, *et al.*, "Metaform optics: bridging nanophotonics and freeform optics," Sci. Adv. **7**, eabe5112 (2021).
203. A. Kirmani, D. Venkatraman, D. Shin, *et al.*, "First-photon imaging," Science **343**, 58–61 (2014).
204. Y. Zhang, Z. He, X. Tong, *et al.*, "Quantum imaging of biological organisms through spatial and polarization entanglement," Sci. Adv. **10**, eadk1495 (2024).

2334-2536/25/010130-27